

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2005, 19:23:59 ; Search time 245 Seconds
(without alignments)
3996.308 Million cell updates/sec

Title: US-10-772-636-64
Perfect score: 10042
Sequence: 1 MVHVARLLLLLTLFLRTDA.....YQFSYRAALEYLGSDHYAT 1912

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10042	100.0	1912	1 PTPD_HUMAN	P23468 homo sapien
2	9280.5	92.4	1894	2 Q64487	Q64487 mus musculus
3	8991.5	89.5	1896	2 Q91AJ1	Q91AJ1 xenopus lae
4	7386.5	73.6	1948	1 PTNS_HUMAN	Q13332 homo sapien
5	7110	70.8	1904	2 Q64699	Q64699 mus musculus
6	7032	70.0	1897	1 PTPF_HUMAN	P10586 homo sapien
7	7004.5	69.8	1898	2 Q86W50	Q86W50 homo sapien
8	6985.5	69.6	1898	2 Q9EQ17	Q9EQ17 mus musculus
9	6981.5	69.5	1898	2 Q64604	Q64604 r protein-t
10	6981	69.5	1887	2 Q9QW67	Q9QW67 rattus sp.
11	6977.5	69.5	1863	2 Q64605	Q64605 rattus norv
12	6750.5	67.2	1788	2 Q91AJ0	Q91AJ0 xenopus lae
13	6308.5	62.8	1254	2 Q8VBV0	Q8VBV0 mus musculus
14	5871.5	58.5	1499	2 Q90M81	Q90M81 homo sapien
15	5794.5	57.7	1502	2 Q7TM17	Q7TM17 mus musculus
16	5748.5	57.2	1501	2 Q9QW00	Q9QW00 rattus sp.
17	5745.5	57.2	1501	2 Q9QW00	Q9QW00 rattus sp.
18	5640	56.2	1529	2 Q8EG86	Q8EG86 mus musculus
19	4935.5	49.1	1889	2 Q7Q0X2	Q7Q0X2 anopheles g
20	4879	48.6	2029	2 Q9V1S8	Q9V1S8 drosophila
21	4877	48.6	2029	1 LAR_DROME	P16621 drosophila
22	4594.5	45.8	1191	2 Q7Z3X4	Q7Z3X4 homo sapien
23	4127	41.1	2051	2 Q44328	Q44328 hirudo medi
24	3942	39.3	749	2 Q8R169	Q8R169 mus musculus
25	3822	38.1	730	2 Q8WX65	Q8WX65 homo sapien
26	3785	37.7	1597	2 Q960M3	Q960M3 drosophila
27	3541	35.3	857	2 Q90YJ4	Q90YJ4 brachydanio
28	3372	33.6	1437	2 Q44329	Q44329 hirudo medi
29	3359	33.4	2200	1 LAR_CAEL	Q9BMN8 caenorhabdi
30	3044	30.3	615	2 Q91A18	Q91A18 xenopus lae
31	2996	29.8	1231	2 Q17024	Q17024 anopheles g

32	2820	28.1	582	2	Q64696	Q64696 mus musculus
33	2504	24.9	508	2	Q90YJ5	Q90YJ5 brachydanio
34	2404	23.9	468	2	Q91BA2	Q91BA2 potamotrygo
35	2381	23.7	468	2	Q91BA0	Q91BA0 potamotrygo
36	2339	23.3	469	2	Q9NL11	Q9NL11 branchiost
37	2304	22.9	468	2	Q91BA5	Q91BA5 potamotrygo
38	2242	22.3	460	2	Q82917	Q82917 rattus norv
39	2226.5	22.2	469	2	Q9NL08	Q9NL08 eptatretus
40	2093	20.8	468	2	Q9NL06	Q9NL06 eptatretus
41	2045.5	20.4	1216	2	Q81120	Q81120 caenorhabdi
42	1954	19.5	398	2	Q62604	Q62604 rattus norv
43	1676	16.7	383	2	Q8MTN0	Q8MTN0 culicoides
44	1651	16.4	472	2	Q9NL02	Q9NL02 eptatretus
45	1586	15.8	793	2	Q91V35	Q91V35 m protein t

ALIGNMENTS

RESULT 1

ID	PTPD_HUMAN	STANDARD;	PRT;	1912 AA.
AC	P23468;			
DT	01-NOV-1991 (Rel. 20, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	Protein-tyrosine phosphatase delta precursor (EC 3.1.3.48) (R-PTP-delta).			
GN	Name=PTPRD;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND MUTAGENESIS OF ARG-1178.			
RP	MEDLINE=95204468; PubMed=7896816; DOI=10.1074/jbc.270.12.6722;			
RA	Pulido R., Krueger N.X., Serra-Pages C., Saito H., Streuli M.;			
RT	"Molecular characterization of the human transmembrane protein-tyrosine phosphatase delta. Evidence for tissue-specific expression of alternative human transmembrane protein-tyrosine phosphatase delta isoforms.";			
RT	J. Biol. Chem. 270:6722-6728(1995).			
RN	[2]			
RP	SEQUENCE OF 390-1912 FROM N.A.			
RC	TISSUE=Placenta;			
RA	MEDLINE=91006018; PubMed=2170109;			
RA	Krueger N.X., Streuli M., Saito H.;			
RT	"Structural diversity and evolution of human receptor-like protein tyrosine phosphatases.";			
RL	EMBO J. 9:3241-3252(1990).			
RN	[3]			
RP	INTERACTIONS WITH PPPIA1; PPPIA2 AND PPPIA3.			
RP	MEDLINE=98288299; PubMed=9624153; DOI=10.1074/jbc.273.25.15611;			
RA	Serra-Pages C., Medley O.G., Tang M., Hart A., Streuli M.;			
RT	"Liprins, a family of LAR transmembrane protein-tyrosine phosphatase-interacting proteins.";			
RL	J. Biol. Chem. 273:15611-15620(1998).			
RN	[4]			
RP	SPICE ISOFORM(S) THAT ARE POTENTIAL NMD TARGET(S).			
RP	PubMed=14759258; DOI=10.1186/gb-2004-5-2-r8;			
RA	Hillman R.T., Green R.E., Brenner S.E.;			
RT	"An unappreciated role for RNA surveillance.";			
RL	Genome Biol. 5:RESEARCH008.1-RESEARCH008.16(2004).			
CC	-1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein tyrosine + phosphate.			
CC	-1- SUBUNIT: Interacts with PPPIA1, PPPIA2 and PPPIA3.			
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	-1- ALTERNATIVE PRODUCTS:			
CC	Event-Alternative splicing; Named isoforms=3;			
CC	Comment-Additional isoforms seem to exist;			
CC	Name=1;			
CC	Isoid=23468-1; Sequence=Displayed;			
CC	Note=May be produced at very low levels due to a premature stop			

CC codon in the mRNA, leading to nonsense-mediated mRNA decay;
CC Name=2; Synonym=Kidney;
CC IsoId=23468-2; Sequence=VSP_005147, VSP_005148, VSP_005149;
CC Name=3; Synonym=Fetal brain;
CC IsoId=23468-3; Sequence=VSP_005150;
CC -1- PPM: A cleavage occurs that separates the extracellular domain
CC from the transmembrane segment.
CC -1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
CC Receptor class 2A subfamily.
CC -1- SIMILARITY: Contains 8 fibronectin type III domains.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 2 protein-tyrosine phosphatase domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL: L38929; AAC1749.1; -;
CC EMBL: X54133; CAA38068.1; -;
CC PIR: A56178; A56178.
CC HSP: P10586; ILAR.
CC Genew: HGNC:9668; PTPRD.
CC MIM: 601598; -;
CC GO: GO:0005887; C: integral to plasma membrane; TAS.
CC GO: GO:0005001; P: transmembrane receptor protein tyrosine pho. . .; TAS.
CC GO: GO:0006470; P: protein amino acid dephosphorylation; TAS.
CC GO: GO:0007195; P: transmembrane receptor protein tyrosine pho. . .; TAS.
CC InterPro: IPR003961; FN III.
CC InterPro: IPR008957; FN III-like.
CC InterPro: IPR003962; FN III subd.
CC InterPro: IPR007110; IG-like.
CC InterPro: IPR003598; IG c2.
CC InterPro: IPR000387; TYR phosphatase.
CC InterPro: IPR000242; Tyr_PP.
CC Pfam: PF00041; fn3; 8.
CC Pfam: PF00047; ig; 3.
CC Pfam: PF00102; Y_phosphatase; 2.
CC PRINTS: PR00014; FNTYPEIII.
CC PRINTS: PR00700; PTPPHPTASE.
CC SMART: SM00060; FN3; 8.
CC SMART: SM00408; IGC2; 3.
CC SMART: SM00194; PTPC; 2.
CC PROSITE: PS00853; IG LIKE; 3.
CC PROSITE: PS00835; FN3; 8.
CC PROSITE: PS00383; TYR_PHOSPHATASE_1; 2.
CC PROSITE: PS00056; TYR_PHOSPHATASE_2; 2.
CC PROSITE: PS00055; TYR_PHOSPHATASE_PTP; 2.
CC Alternative splicing: Glycoprotein; Hydrolase; Immunoglobulin domain;
CC Protein phosphatase; Receptor; Repeat; Signal; Transmembrane.
CC SIGNAL 1 20
CC CHAIN 1 1912 Protein-tyrosine phosphatase delta.
CC DOMAIN 21 1265 Extracellular (Potential).
CC TRANSMEM 1266 1290 Potential.
CC DOMAIN 1291 1912 Cytoplasmic (Potential).
CC DOMAIN 24 114 Ig-like C2-type 1.
CC DOMAIN 126 224 Ig-like C2-type 2.
CC DOMAIN 236 318 Ig-like C2-type 3.
CC DOMAIN 323 411 Fibronectin type-III 1.
CC DOMAIN 417 511 Fibronectin type-III 2.
CC DOMAIN 516 604 Fibronectin type-III 3.
CC DOMAIN 609 706 Fibronectin type-III 4.
CC DOMAIN 711 819 Fibronectin type-III 5.
CC DOMAIN 824 913 Fibronectin type-III 6.
CC DOMAIN 918 1013 Fibronectin type-III 7.
CC DOMAIN 1017 1103 Fibronectin type-III 8.
CC DOMAIN 1375 1618 Protein-tyrosine phosphatase 1.
CC DOMAIN 1619 1912 Protein-tyrosine phosphatase 2.
CC ACT_SITE 1553 1553 Phosphotyrosine intermediate (By similarity).

FT	ACT_SITE	1844	1844	1844	Phosphotyrosine intermediate (By similarity)
FT	SITE	1175	1178	1178	Cleavage (Potential).
FT	CARBOHYD	254	254	254	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	299	299	299	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	724	724	724	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	832	832	832	N-linked (GlcNAc. . .) (Potential).
FT	VARSPLIC	181	189	189	Missing (in isoform 2). /FTId=VSP_005147.
FT	VARSPLIC	226	229	229	Missing (in isoform 2). /FTId=VSP_005148.
FT	VARSPLIC	775	783	783	Missing (in isoform 2). /FTId=VSP_005149.
FT	VARSPLIC	609	1137	1137	Missing (in isoform 3). /FTId=VSP_005150.
FT	MUTAGEN	1178	1178	1178	R->A: 2.5-fold reduction in cleavage.
FT	SEQUENCE	1912	AA; 214759	MM; 3A88CBDCD32182E26	CRC64;
SQ	SEQUENCE	1912	AA; 214759	MM; 3A88CBDCD32182E26	CRC64;
Query Match 100.0%; Score 10042; DB 1; Length 1912;					
Best Local Similarity 100.0%; Pred. No. 0;					
Matches 1912; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	1	MVHVARELLLLLTFFLRDAETPPRTTRTPVDQTVSGGVSASFICQATGDPKIVNNKK	60		
Db	1	MVHVARELLLLLTFFLRDAETPPRTTRTPVDQTVSGGVSASFICQATGDPKIVNNKK	60		
Qy	61	GKVSQRFVEIEFDDGGSGVLRIQPLTRPRDAIEYECVSNVNGEISVSTRITVLREDQ	120		
Db	61	GKVSQRFVEIEFDDGGSGVLRIQPLTRPRDAIEYECVSNVNGEISVSTRITVLREDQ	120		
Qy	121	IPRGFTIDMGPKLVVERTRTATMLCAASGNPDPEITWPKDFLPVDTNNNGRIKQLRS	180		
Db	121	IPRGFTIDMGPKLVVERTRTATMLCAASGNPDPEITWPKDFLPVDTNNNGRIKQLRS	180		
Qy	181	ESIGGTPIRGALQIEQSESDQGYECVATNSAGTRYASAPANLYVRELREVRVPPPSI	240		
Db	181	ESIGGTPIRGALQIEQSESDQGYECVATNSAGTRYASAPANLYVRELREVRVPPPSI	240		
Qy	241	PPTNHEIMPGGSVNITCVAVGSPMPYKWLGAEDLTPEDDMPGRNVLENDVRSQANY	300		
Db	241	PPTNHEIMPGGSVNITCVAVGSPMPYKWLGAEDLTPEDDMPGRNVLENDVRSQANY	300		
Qy	301	TCVAMSTGLVIEALQITVAKLPKPGTPVVTSTSTITLTWDSGNPEPVSYYIIQHKP	360		
Db	301	TCVAMSTGLVIEALQITVAKLPKPGTPVVTSTSTITLTWDSGNPEPVSYYIIQHKP	360		
Qy	361	KNSEELYKEIDGVATTRYSVAGLSPYDIEPRVAVVANNIGRPPSEPLVTQTSQAPSSA	420		
Db	361	KNSEELYKEIDGVATTRYSVAGLSPYDIEPRVAVVANNIGRPPSEPLVTQTSQAPSSA	420		
Qy	421	PRDVOARMLSTTILVQWKEPEENGQIQGYRVVYTMPTDPTOHVNNMKNHVNADSOITIG	480		
Db	421	PRDVOARMLSTTILVQWKEPEENGQIQGYRVVYTMPTDPTOHVNNMKNHVNADSOITIG	480		
Qy	481	NLVQPKTYSVKVLAFPTSGDGLSSDIQVITQTGVPGQPLNFKAPESETSILLISWTTPR	540		
Db	481	NLVQPKTYSVKVLAFPTSGDGLSSDIQVITQTGVPGQPLNFKAPESETSILLISWTTPR	540		
Qy	541	SDTIANYELVYKDGHEGEQRITTEPGTSYRLOGKPNLSYFYRLAARSPQGLGASTAEI	600		
Db	541	SDTIANYELVYKDGHEGEQRITTEPGTSYRLOGKPNLSYFYRLAARSPQGLGASTAEI	600		
Qy	601	SARTMQSKPSAPPQDICTSPSSSTILVSWQPPVPEKONGIITEYSIKYTAVDGDDKPH	660		
Db	601	SARTMQSKPSAPPQDICTSPSSSTILVSWQPPVPEKONGIITEYSIKYTAVDGDDKPH	660		
Qy	661	EILGIPSDTTKYLEQLEKTEYRITVTAHTDVGPGESLSVLRTNEDVSPGPKVEV	720		
Db	661	EILGIPSDTTKYLEQLEKTEYRITVTAHTDVGPGESLSVLRTNEDVSPGPKVEV	720		
Qy	721	EAVNSTSVKSWRSPVNNKQGRGQVHVYVRMENGEPKQPMKQVMLADAQWEPDDT	780		
Db	721	EAVNSTSVKSWRSPVNNKQGRGQVHVYVRMENGEPKQPMKQVMLADAQWEPDDT	780		

QY 781 TEHDMIIISGLQETSYSLTVTATYTTKDGARSKPKLVSTTGAVPGKPRLVINHTQNTAL 840
DB 781 TEHDMIIISGLQETSYSLTVTATYTTKDGARSKPKLVSTTGAVPGKPRLVINHTQNTAL 840
QY 841 IQWHPPVDTFGPLOGVRLKFGKMEPLTTLFSEKEDHFTATDTHKGASVYFRLSARNK 900
DB 841 IQWHPPVDTFGPLOGVRLKFGKMEPLTTLFSEKEDHFTATDTHKGASVYFRLSARNK 900
QY 901 VGFGEWVKBIISPEVPTGFPQNLHSECTTSTVQLSWQPPVLAERNGIITKYLLYRD 960
DB 901 VGFGEWVKBIISPEVPTGFPQNLHSECTTSTVQLSWQPPVLAERNGIITKYLLYRD 960
QY 961 INIPLLPHEQLIVPADDTTMTLGLKPDYTDYKVRHAHSGKPGTSPSVQFTPLVDQVF 1020
DB 961 INIPLLPHEQLIVPADDTTMTLGLKPDYTDYKVRHAHSGKPGTSPSVQFTPLVDQVF 1020
QY 1021 AKNFHVKAAMKTSVLLSWEIPEYNASAMPFKLYDDGKVVVEVDGRATQKLLVNLKPKS 1080
DB 1021 AKNFHVKAAMKTSVLLSWEIPEYNASAMPFKLYDDGKVVVEVDGRATQKLLVNLKPKS 1080
QY 1081 YSFVLTRNGNSAGGLQHRVTAKTAPDVLRTKPAFIGNKLDGMITVQLPEVPANENIKGY 1140
DB 1081 YSFVLTRNGNSAGGLQHRVTAKTAPDVLRTKPAFIGNKLDGMITVQLPEVPANENIKGY 1140
QY 1141 YIIIVPLKSRKFKIPKWPESPEMELDELLKESIKRRSIRYGRVVELKPYIAAHFDVLP 1200
DB 1141 YIIIVPLKSRKFKIPKWPESPEMELDELLKESIKRRSIRYGRVVELKPYIAAHFDVLP 1200
QY 1201 TEFTLGDDKHGFGFTNKLOSQOEYVFFVLAHMEHAESKMYATSPYDPVMSMDLDPOPI 1260
DB 1201 TEFTLGDDKHGFGFTNKLOSQOEYVFFVLAHMEHAESKMYATSPYDPVMSMDLDPOPI 1260
QY 1261 TDEEGLIWWGVPVLAIVFIIICIVIAILLYKRAESDSRKSIIINNKEIPSHHPTDVE 1320
DB 1261 TDEEGLIWWGVPVLAIVFIIICIVIAILLYKRAESDSRKSIIINNKEIPSHHPTDVE 1320
QY 1321 LRLNFTQPGWASHPPPILELAHIERLKANDNLKFSQOYESIDPGQOFTWEHNSLEVN 1380
DB 1321 LRLNFTQPGWASHPPPILELAHIERLKANDNLKFSQOYESIDPGQOFTWEHNSLEVN 1380
QY 1381 KPKRYANVIAVDHRSVLLSAIEGIPGSDYNNANVYIDGVRKQNAVIAIOGSLPETFGDFW 1440
DB 1381 KPKRYANVIAVDHRSVLLSAIEGIPGSDYNNANVYIDGVRKQNAVIAIOGSLPETFGDFW 1440
QY 1441 RMWIEQRSAIVMTKLEERSRVKCDQYWPSSRGTEHGLVQVTLDDTVELATYCVRTFAL 1500
DB 1441 RMWIEQRSAIVMTKLEERSRVKCDQYWPSSRGTEHGLVQVTLDDTVELATYCVRTFAL 1500
QY 1501 YKNGSSEKREVRQFTAMPDGHGVEHPTPLFLRVRKTCNPPDAGPMVHCSAGVGT 1560
DB 1501 YKNGSSEKREVRQFTAMPDGHGVEHPTPLFLRVRKTCNPPDAGPMVHCSAGVGT 1560
QY 1561 GCFVIDAMLERIKHEKTVDIYGHVTLMAQRNVYVQTEDOYIFTHDALLAEAVTCGNTVEV 1620
DB 1561 GCFVIDAMLERIKHEKTVDIYGHVTLMAQRNVYVQTEDOYIFTHDALLAEAVTCGNTVEV 1620
QY 1621 PARNLYAYIKLQIETGENTVGMLEFKRLASSKAHTSRFISANLPCNKFKNRLVNIIMP 1680
DB 1621 PARNLYAYIKLQIETGENTVGMLEFKRLASSKAHTSRFISANLPCNKFKNRLVNIIMP 1680
QY 1681 YESTRVCLOPIRGVSGSYINASFIDGVRQKAYIATOGPLAETTEDFWMLWEHNSIV 1740
DB 1681 YESTRVCLOPIRGVSGSYINASFIDGVRQKAYIATOGPLAETTEDFWMLWEHNSIV 1740
QY 1741 VMLTKLRMGREKCHOYPAERSARYQYFVVDPMAYNNMPOYLREFKVTVDARQOSRTV 1800
DB 1741 VMLTKLRMGREKCHOYPAERSARYQYFVVDPMAYNNMPOYLREFKVTVDARQOSRTV 1800
QY 1801 RQFQTDWPEQGVKPSGSGFIDFQGVHKTKEQFGQDGPISVHCSAGVGTGVFTLSIV 1860
DB 1801 RQFQTDWPEQGVKPSGSGFIDFQGVHKTKEQFGQDGPISVHCSAGVGTGVFTLSIV 1860

QY 1861 LERMYEGVVDIFQTVKMLRTQRPAMVQTEDOYQFSYRAALYLGSDHYAT 1912
DB 1861 LERMYEGVVDIFQTVKMLRTQRPAMVQTEDOYQFSYRAALYLGSDHYAT 1912
RESULT 2
Q84487
ID Q64487 PRELIMINARY; PRT; 1894 AA.
AC Q64487; Q64486; Q64488; Q64495;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Protein-tyrosine phosphatase, receptor-type, D precursor (EC 3.1.3.48)
DE (Protein-tyrosine phosphatase delta) (R-PTP-delta).
GN Name=Ptpcd;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=BRAIN;
RX MEDLINE=93360986; PubMed=8355697;
RA Mizuno K., Hasegawa K., Katagiri T., Ogimoto M., Ichikawa T.,
RA Yakura H.;
RT "MPTP delta, a putative murine homolog of HPTP delta, is expressed in
RT specialized regions of the brain and in the B-cell lineage.";
RL Mol. Cell. Biol. 13:5513-5523(1993).
RN [2]
SEQUENCE OF 1430-1534 FROM N.A.
RC STRAIN=BALB/C; TISSUE=BRAIN;
RX MEDLINE=95134232; PubMed=7832766;
RA Hendriks W., Schepens J., Brugman C., Zeeuwen P., Wieringa B.;
RT "A novel receptor-type protein tyrosine phosphatase with a single
RT catalytic domain is specifically expressed in mouse brain.";
RL Biochem. J. 305:499-504(1995).
CC -|- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O = PROTEIN
CC TYROSINE + PHOSPHATE.
CC -|- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -|- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Comment=Additional isoforms seem to exist;
CC Name=C;
CC IsoId=Q64487-1; Sequence=Displayed;
CC Name=A;
CC IsoId=Q64487-2; Sequence=VSP_050406, VSP_050407, VSP_050408;
CC Name=B;
CC IsoId=Q64487-3; Sequence=VSP_050406, VSP_050407;
CC -|- TISSUE SPECIFICITY: BRAIN, KIDNEY, HEART, AND SOME B-CELL LINES.
CC -|- PTM: A CLEAVAGE OCCURS THAT SEPARATES THE EXTRACELLULAR DOMAIN
CC FROM THE TRANSMEMBRANE SEGMENT (BY SIMILARITY).
CC -|- SIMILARITY: EXTRACELLULAR REGION TYPICAL OF A CAM FAMILY (3 IG-
CC LIKE DOMAINS FOLLOWED BY 8 FIBRONECTIN TYPE III-LIKE DOMAINS), AND
CC A CYTOPLASMIC REGION COMPOSED OF TWO PTASE DOMAINS.
EMBL; D13903; BAA03003.1; -;
EMBL; D13905; BAA03005.1; -;
EMBL; D13904; BAA03004.1; -;
EMBL; Z23051; CAA80586.1; -;
PIR; C54689; C54689.
PIR; D54689; D54689.
HSP; P10586; 1LAR.
DR GO:0016031; C: integral to membrane; IEA.
DR GO:0016787; F: hydrolase activity; IEA.
DR GO:0004725; F: protein tyrosine phosphatase activity; IEA.
DR GO:0004872; F: receptor activity; IEA.
DR GO:0006470; P: protein amino acid dephosphorylation; IEA.
DR InterPro; IPR003962; FNIII subd.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; Iq c2.
DR InterPro; IPR000387; TYR phosphatase.
DR InterPro; IPR000242; Tyr_Pp.

DR	Pfam; PF00041; fn3; 7.	
DR	Pfam; PF00047; ig; 1.	
DR	Pfam; PF00102; y_phosphatase; 2.	
DR	PRINTS; PR00102; ENTYPBIII.	
DR	PRINTS; PR00700; PRTYPHPTASE.	
DR	SMART; SM00060; FN3; 8.	
DR	SMART; SM00408; IGC2; 3.	
DR	SMART; SM00194; PTPC; 2.	
DR	PROSITE; PS00853; FN3; 8.	
DR	PROSITE; PS00835; IG LIKE; 3.	
DR	PROSITE; PS00383; TYR_PHOSPHATASE 1; 2.	
DR	PROSITE; PS00056; TYR_PHOSPHATASE 2; 2.	
DR	PROSITE; PS00055; TYR_PHOSPHATASE_PTP; 2.	
Kw	Alternative splicing; Glycoprotein; Hydrolase; Immunoglobulin domain;	
Kw	Receptor; Repeat; Signal; Transmembrane.	
FT	SIGNAL 1 18	
FT	CHAIN 19 1894	
FT	DOMAIN 19 1251	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 1252 1272	POTENTIAL.
FT	DOMAIN 1273 1894	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN 45 112	IG-LIKE C2-TYPE DOMAIN.
FT	DOMAIN 147 215	IG-LIKE C2-TYPE DOMAIN.
FT	DOMAIN 245 304	IG-LIKE C2-TYPE DOMAIN.
FT	DOMAIN 315 409	FIBRONECTIN TYPE-III.
FT	DOMAIN 412 508	FIBRONECTIN TYPE-III.
FT	DOMAIN 511 601	FIBRONECTIN TYPE-III.
FT	DOMAIN 603 702	FIBRONECTIN TYPE-III.
FT	DOMAIN 705 806	FIBRONECTIN TYPE-III.
FT	DOMAIN 809 899	FIBRONECTIN TYPE-III.
FT	DOMAIN 901 1000	FIBRONECTIN TYPE-III.
FT	DOMAIN 1004 1122	FIBRONECTIN TYPE-III.
FT	DOMAIN 1360 1601	PROTEIN-TYROSINE PHOSPHATASE.
FT	DOMAIN 1602 1894	PROTEIN-TYROSINE PHOSPHATASE.
FT	DISULFD 52 105	POTENTIAL.
FT	DISULFD 154 208	POTENTIAL.
FT	DISULFD 252 297	POTENTIAL.
FT	ACT SITE 1536	BY SIMILARITY.
FT	ACT SITE 1826	BY SIMILARITY.
FT	SITE 1160 1163	CLEAVAGE (POTENTIAL).
FT	VARSPLIC 1 1	M -> MCLTSCFILASHMLSCDLVFP (in isoform A and isoform B).
FT		/FTId=VSP 050406.
FT	VARSPLIC 2 225	Missing (in isoform A and isoform B).
FT		/FTId=VSP 050407.
FT	VARSPLIC 602 1001	Missing (in isoform A).
FT		/FTId=VSP 050408.
FT	CONFLICT 1430 1432	EQE -> QRS (IN REF. 2).
FT	CONFLICT 1475 1485	HILCPDICTIN -> ATYCVRTFALYK (IN REF. 2).
FT	CONFLICT 1493 1493	K -> E (IN REF. 2).
FT	CONFLICT 1514 1516	VPS -> FLA (IN REF. 2).
FT	CONFLICT 1519 1525	TESQNLH -> RRVKTCN (IN REF. 2).
FT	SEQUENCE 1894 AA; 212194 MW; 8211033B53DD163C CRC64;	
Qy	Query Match	92.4%; Score 9280.5; DB 2; Length 1894;
Qy	Best Local Similarity	93.8%; Pred. No. 0;
Qy	Matches 1796; Conservative 31; Mismatches 57; Indels 31; Gaps 12;	
Qy	1 MYHVARLLLLLTFFLRDTDAETPPRFTPTVDQTVSGGVASFICQATGDRPKIVWNKK 60	
Qy	8 MYPVRPLSLLTFFLCACAEPTPTPTVDQTVSGGVASFICQATGDRPKIVWNKK 67	
Qy	61 GKXVSNQRFVEJEPDGGSVLRQLPRTPRDEAIEYCVASNNVGEISVSTRLTVLREDQ 120	
Qy	68 GKXVSNQRFVEJEPDGGSVLRQLPRTPRDEAIEYCVASNNVGEISVSTRLTVLREDQ 127	
Qy	121 PRGFPTIDMGQPKVVERTRATMLCAASGNPDPEITWFKDFLPVDTNNNGRIKQLRS 180	
Qy	128 PRGFPTIDMGQPKVVERTRATMLCAASGNPDPEITWFKDFLPVDTNNNGRIKQLRS 187	
Qy	181 ESIGGTPIRGALQIQSEESDQGYECVATNSAGTRYAPANLYVRELVRVPPRFSI 240	
Qy	188 ESI-----GALQIQSEESDQGYECVATNSAGTRYAPANLY-----VRRVPPRFSI 235	

Qy	241 PPTNHEIMPGGSVNITCAVGSMPYKWMGAEDLTPEDDMPIGRNVLELNDVROSANY 300	
Qy	236 PPTNHEIMPGGSVNITCAVGSMPYKWMGAEDLTPEDDMPIGRNVLELNDVROSANY 295	
Qy	301 TCVMSTLGVIEAIAQITVKALPKPCTPVVTESTATSIITLWDSGNPEPVSYYIIQHKP 360	
Qy	296 TCVMSTLGVIEAIAQITVKALPKPCTPVVTESTATSIITLWDSGNPEPVSYYIIQHKP 355	
Qy	361 KNSSELYKEIDGVATTRYSVAGLSYSDYBFRVAVANNIGRPPSEPVLTQTSBOAPSSA 420	
Qy	356 KNSSELYKEIDGVATTRYSVAGLSYSDYBFRVAVANNIGRPPSEPVLTQTSBOAPSSA 415	
Qy	421 PRDQVARMLSSTLILVQWKEPEEPNGQIQYRVYVYTDPTQHVNNWKNHVNADSOIITIG 480	
Qy	416 PRDQVARMLSSTLILVQWKEPEEPNGQIQYRVYVYTDPTQHVNNWKNHVNADSOIITIG 475	
Qy	481 NLVPQKTSYKVLAFSTIGDPLSSDIQVITQTCVPGQPLNFKAPESETSILLSWTPPR 540	
Qy	476 NLVPQKTSYKVLAFSTIGDPLSSDIQVITQTCVPGQPLNFKAPESETSILLSWTPPR 535	
Qy	541 SDTIANYELVYKDGEGHEEQRIITIEPCTSVRLQGLKPNLSLYYFRLAARSPQGLGASTAEI 600	
Qy	536 SDTIANYELVYKDGEGHEEQRIITIEPCTSVRLQGLKPNLSLYYFRLAARSPQGLGASTAEI 595	
Qy	601 SARTMQSKSPAPPODISCTSPSSSTILVSWQPPVEKQNGIITEYSIKYTAVDGEDDKPH 660	
Qy	596 SARTMQSKSPAPPODISCTSPSSSTILVSWQPPVEKQNGIITEYSIKYTAVDGEDDKPH 654	
Qy	661 EILGIPSDTTKYLLQLEQEKWTEYRITVTAHTDVGPGPESISVLIRTNEDVSPGPRKVEV 720	
Qy	655 EILGIPSDTTKYLLQLEQEKWTEYRITVTAHTDVGPGPESISVLIRTNEDVSPGPRKVEV 714	
Qy	721 EAVNSTSVKVSWSRSPVKNQHGQIRGYQVHYVRMENGEPKQPMKQVMLADADAQWFFDDT 780	
Qy	715 EAVNSTSVKVSWSRSPVKNQHGQIRGYQVHYVRMENGEPK-SAMLKDVMLADAQ----- 767	
Qy	781 TEHMIISGLQPEPSTYSITVYTAHTKDGARSKPLVSTTGAVPGKRLVNHQMTAL 840	
Qy	768 ---DMIISGLQPEPSTYSITVYTAHTKDGARSKPLVSTTGAVPGKRLVNHQMTAL 824	
Qy	841 IQWHPPVDVTFGLQGYRLKFKGRKDMPELTTLEFSEKEDHFTATDIHKGASVYFRLSARNK 900	
Qy	825 IQWHPPVDVTFGLQGYRLKFKGRKDMPELTTLEFSEKEDHFTATDIHKGASVYFRLSARNK 883	
Qy	901 VGFGEEMVKEISIEEVPPTGPQNLHSEGTSTSTSVQLSWQPPVLAERNGIITKYLTLVYRD 960	
Qy	884 VGFGEEMVKEISIEEVPPTGPQNLHSEGTSTSTSVQLSWQPPVLAERNGIITKYLTLVYRD 943	
Qy	961 INIPLPNEQLIVPADTMTLTGLKPDYDVKVRHTSKGPGVSPSPVQFRTLVPDQ-V 1019	
Qy	944 INVPLPMEHLIVPADTMTLTGLKPDYDVKVRHTSKGPGVSPSPVQFRTLVPDQSM 1003	
Qy	1020 FAKHFKVAVMKTSVLLSWEIPENYNAMPFK-ILYDDGKMVEEVDGRATQKLI.VNLKPE 1078	
Qy	1004 FAKHFKVAVMKTSVLLSWEIPENYNAMPFK-ILYDDGKMVEEVDGRATQKLI.VNLKPE 1063	
Qy	1079 KYSFVLITNRNSAGGLQHRVTAKTAPDLVTKTAPAFIKTNLDGMITVQLDEVPAENIK 1138	
Qy	1064 KYSFVLITNRNSAGGLQHRVTAKTAPDLVTKTAPAFIKTNLDGMITVQLDEVPAENIK 1123	
Qy	1139 GYVLIIVPLKKSCKFTKPWESPDMEIDELLKEISRKRRSIRYGREVELKPYIAAHFDV 1198	
Qy	1124 GYVLIIVPLKKSCKFTKPWESPDMEIDELLKEISRKRRSIRYGREVELKPYIAAHFDV 1183	
Qy	1199 LPTEFTLGDGKHGGFTNKQLQSGQYVFFVLAVMHAESKMYATSPSPVSMDLDPQ 1258	
Qy	1184 LPTEFTLGDGKHGGFTNKQLQSGQYVFFVLAVMHAESKMYATSPSPVSMDLDPQ 1243	
Qy	1259 PITDEEGLIIVWGVPLAVWPIICIVIAILLIKRKRAESDSKRSIPNNKEIPSHHPTDP 1318	
Qy	1244 PITDEEGLIIVWGVPLAVWPIICIVIAILLIKRKRAESDSKRSIPNNKEIPSHHPTDP 1303	

QY 1319 VELRLNFTPCWASHPPPILELADHIERLKANDNLKFSQYESIDPGQQFTWHSNLE 1378
 Db 1304 VELRLNFTPCWASHPPPILELADHIERLKANDNLKFSQYESIDPGQQFTWHSNLE 1363
 QY 1379 VNKPNRNVANVAYDHSRVLSSAIEGIPGSDVYVANYIDGKQKAYATQGSLETETGD 1438
 Db 1364 VNKPNRNVANVAYDHSRVLSSAIEGIPGSDVYVANYIDGKQKAYATQGSLETETGD 1423
 QY 1439 FWRMIWORSATVVMWTKLEERSVKDCQWPSRGTEHGLVQVTLTLLDTVELATYCVRTF 1498
 Db 1424 FWRMIWEO- EATVVMWTKLEERSVKDCQWPSRGTEHGLVQVTLTLLDTVELHILCPDIC 1482
 QY 1499 ALYKNGSSEKREVRQFOTAWPDHGVPEHPTFLAFLRVKTCNPDAGPMVHCSAGVG 1558
 Db 1483 TL- NNGSSSEKRVQFOTAWPDHGVPEHPTFLAFLRVKTCNPDAGPMVHCSAGVG 1541
 QY 1559 RTGCFIVIDAMLERIKHEKTVDIYGHVTLMAQRNMVQTEQYIFIHDALEAVTCGNT 1618
 Db 1542 RTGCFIVIDAMLERIKHEKTVDIYGHVTLMAQRNMVQTEQYIFIHDALEAVTCGNT 1601
 QY 1619 EVPARNLVAYIQKLTQIETGENVTGMELEFKRLASSKAHTSRFISANLPCNKFQRLVNI 1678
 Db 1602 EVPARNLVAYIQKLTQIETGENVTGMELEFKRLASSKAHTSRFISANLPCNKFQRLVNI 1661
 QY 1679 MPYESTRVCLOPIRGVSGSDYINASFIDGROOKAYIATQGLAETTEDFWRMLWEHNS 1738
 Db 1662 MPYESTRVCLOPIRGVSGSDYINASFIDGROOKAYIATQGLAETTEDFWRMLWEHNS 1721
 QY 1739 IVWMLTKLRMGREKCHOYPAERSARYQYFVDPMAEYNNPQYILREFKVTDARDG-QS 1797
 Db 1722 IVWMLTKLRMGREKCHOYPAERSARYQYFVDPMAEYNNPQYILREFKVTDARDG-QS 1779
 QY 1798 RTVRFQFTDMPNQVPSGSGFIDFQGVHKTKEQFGDQSPISVHCSAGVGTGVFTL 1857
 Db 1780 RTVRFQFTDMPNQVPSGSGFIDFQGVHKTKEQFGDQSPISVHCSAGVGTGVFTL 1839
 QY 1858 SILVERMEYGVVDIFQTVKMLRTORPAMVQTEQYQFSYAALEYLGSFDHYAT 1912
 Db 1840 SILVERMEYGVVDIFQTVKMLRTORPAMVQTEQYQFSYAALEYLGSFDHYAT 1894
 RESULT 3
 Q91AJ1 PRELIMINARY; PRT; 1896 AA.
 ID Q91AJ1
 AC Q91AJ1
 DT 01-OCT-2000 (TremBLrel. 15, Created)
 DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
 DE Receptor protein tyrosine phosphatase delta.
 GN Name=XPTP-D;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20193505; PubMed=10727868; DOI=10.1016/S0925-4773(99)00345-7;
 RA Johnson K.G., Holt C.E.;
 RT "Expression of CrYP-alpha, LAR, PTP-delta, and PTP-rho in the
 RT developing Xenopus visual system.";
 RL Mech. Dev. 92:291-294(2000).
 DR EMBL; AF197944; AAF3605.1; -.
 DR HSSP; P10586; ILAR.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR GO; GO:0004725; F:protein tyrosine phosphatase activity; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0006470; F:protein amino acid dephosphorylation; IEA.
 DR InterPro; IPR003962; FNIII_subd.
 DR InterPro; IPR003961; FN III.
 DR InterPro; IPR008957; FN III-like.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003598; IG_C2.

DR InterPro; IPR000387; TYR_phosphatase.
 DR InterPro; IPR000242; Tyr_PP.
 DR Pfam; PF00041; fn3; 8.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00102; Y_phosphatase; 2.
 DR PRINTS; PR00014; ENTYPBIII.
 DR PRINTS; PR00700; PRTYPHPTASE.
 DR SMART; SM00060; FN3; 8.
 DR SMART; SM00408; IGC2; 3.
 DR SMART; SM00194; PTPC; 2.
 DR PROSITE; PS50853; FN3; 8.
 DR PROSITE; PS50835; IG_LIKE; 3.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 2.
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 2.
 DR PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 2.
 KW Hydrolase; Receptor.
 SQ SEQUENCE 1896 AA; 213067 MW; 08AC9003034199A4 CRC64;
 Query Match 89.5%; Score 8991.5; DB 2; Length 1896;
 Best Local Similarity 88.9%; Pred. No. 0;
 Matches 1699; Conservative 101; Mismatches 88; Indels 23; Gaps 4;
 QY 2 VHVARLLLLLTFFLRDTDAETPPRFRTRTPVDTGVSGGVASFIQATGDPKPKIVNKKG 61
 Db 9 MNIARPVVLLCLFLLHAGAETPPKLTPTVDQIGVSGGVASFIQATGDPKPKIVNKKG 68
 QY 62 KKVSNQRFEVIEFDGGGSLVRIQLRTPRDEAIYECVASNNGEISVSTRITVLREDO 121
 Db 69 KKVSNQRFEVIEFDGGGSLVRIQLRTPRDEAIYECVASNNGEISVSTRITVLREDO 128
 QY 122 PRGPTTDMGQPKVETRTATMLCAASGNDPDEITWFKDFLPVDTNNNGRIKQLRSE 181
 Db 129 PRGPTTDMGQPKVETRTATMLCAASGNDPDEITWFKDFLPVDTNNNGRIKQLRSE 187
 QY 182 SIGGTPIRGALQIQSSESDQKYECAVATNSAGTRYSAPANLYVRELREVRVPRPSTIP 241
 Db 188 -----GALQIQSSESDQKYECAVATNSAGTRYSAPANLYVRELREVRVPRPSTIP 235
 QY 242 PTNHEIMPGGSVNITCVAVGSPMPYVWMLGAEDLTPEDDMPIGRNVLNLDVROSANYT 301
 Db 236 PTNHEIMPGGSVNITCVAVGSPMPYVWMLGAEDLTPEDDMPIGRNVLNLDVROSANYT 295
 QY 302 CVANSTGLVIEAIAQITVKALPKPGTPVVTETATSTILTWDGNGNPFVYIIQHKPK 361
 Db 296 CVANSTGLVIEAIAQINVKALPKPGTPVVTETATSTILTWDGNGNPFVYIIQHKPK 355
 QY 362 NSEELYKEIDGVAATRYSVAGLSYDVEFRVAVANNIGRGPPEPVLTTOTSEQAPSSAP 421
 Db 356 SSEEQYKEIDGVAATRYSVAGLSYDVEFRVAVANNIGRGPPEPVLTTOTSEQAPSSAP 415
 QY 422 RDVQARMLSTTILVQWKEPEEPNGQIQGYRVYVYTDPTQHVNNMMKHNVADSQITIGN 481
 Db 416 RNVAQARMLSTTILVQWKEPEEPNGQIQGYRVYVYTDPTQHVNNMMKHNVADSQITIGN 475
 QY 482 LVPQKTSVKVLAFTSISGDGLSSDIQVITQTVGQPLNFKAPESESTILLSWTPPRS 541
 Db 476 LEPQKTSVKVLAFTSISGDGLSSDIQVITQTVGQPLNFKAPESESTILLSWTPPRS 535
 QY 542 DTIANEYLVKDGEGHEEORITIEBPTGSYRLQGLKPNLSLYFRLAARSPOGLGASTAIS 601
 Db 536 DTISSYLYKDGDAHEEV-ITIDPATSYRLQGLKPNLSLYFRLAARSPOGLGASTAIS 594
 QY 602 ARTMQSKPSAPPQDISCTSPSSSTILVSWQPPVPEKQNGIITEYSIKYTAVDGDDKPE 661
 Db 595 ARTMQSKPSAPPQDIRCNSQSSSTILVSWQPPVPEKQNGIITEYSIKYTAVDGDDKPE 654
 QY 662 ILGTPSDTTKYLLEQLEKWTYRITVTAHTDVGCPESLSVLITNEDVPSGPRKVEVE 721
 Db 655 ILGTSSTQYLLLEQLEKWTYRITVTAHTDVGCPESLSVLITNEDVPSGPRKVEVE 714
 QY 722 AVNSTSVKSVNRSVPVKNQHQIRGYVHYVYRMENBPKGQPMKLDVNLADAQWEDDTT 781
 Db 715 AVNSTSVKSVNRSVPVKNQHQIRGYVHYVYRMENBPKGQPMKLDVNLADAQWEDDTT 767

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 5, 2005, 19:59:42 ; Search time 2966 Seconds
(without alignments)
12500.115 Million cell updates/sec

Title: US-10-772-636-63
Perfect score: 6263
Sequence: 1 gctaactcaaggagacgtc.....aagtgaattttaaaaaagtt 6263

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2359870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6263	100.0	6263	11	ADI32147 Human CDN
2	6263	100.0	6263	13	ADR40182 Human pro
3	2473	39.5	6363	10	ADD18743 Human dis
4	2473	39.5	6363	11	ADI31553 Human CDN
5	2398	38.3	6000	2	AAQ86478 Human pTP
6	2398	38.3	6000	2	AAT85389 Human pro
7	2398	38.3	6000	2	AAx06095 Human pro
8	2398	38.3	6000	4	AAx02144 Human pro
9	2374	37.9	5913	12	ADQ84008 Human tum
10	2365.4	37.8	4555	2	AAQ94311 Tyrosine
11	2359	37.7	7741	4	AAx22448 Human CDN
12	2357.4	37.6	7702	3	AAZ91908 Human pro
13	2357.4	37.6	7702	3	AAZ59133 LAR tyros
14	2357.4	37.6	7702	3	AAx88739 Human pro
15	2357.4	37.6	7702	10	ADD18741 Human dis
16	2357.4	37.6	7702	10	ADK61221 Ovarian c
17	2357.4	37.6	7702	11	ADI32010 Human CDN
18	2357.4	37.6	7702	12	ADP18653 Human TAT
19	2357.4	37.6	7702	13	ADJ33475 Human leu
20	2357.4	37.6	7724	13	ADJ33489 Human LAR

21	2357.4	37.6	7945	5	ABV27897	Human pro
c 22	2348	37.5	7705	4	AAH98405	Human EST
c 23	2348	37.5	7705	4	AAx22684	Human CDN
24	2346.6	37.5	6545	10	ADB52922	Primary r
25	2307.8	36.8	6734	6	ABI19344	Mouse iac
26	2303.2	36.8	8232	11	ACN89826	Breast ca
27	2233.6	35.9	7007	12	ADQ85067	Human tum
28	2173.8	34.7	6469	10	ADB58330	Toxicity r
29	2173.8	34.7	6469	10	ADB52889	Primary r
30	1901.8	30.4	6050	9	AAZ57343	Human kin
31	1517.6	24.2	4836	12	ADQ63853	Novel hum
32	1345.8	21.5	4078	10	ACA56534	Human sig
33	1345.8	21.5	4078	12	ADI56330	Human pol
34	1300	20.8	5690	2	AAQ86902	Rat recep
35	1290	20.6	6143	13	ADJ33482	Mouse leu
36	1245.2	19.9	5412	6	ABI19920	Rat mucoc
37	1200	19.2	4600	13	ADR06902	Full leng
38	1196.6	19.1	4190	10	ADB58003	Toxicity-
39	1095.8	17.5	3467	3	AAZ91907	Human pro
40	1095.8	17.5	3467	3	AAZ59132	LAR tyros
41	953	15.2	1200	2	AAQ94313	Tyrosine
42	851.2	13.6	7630	4	ABL30223	Drosophil
43	830.6	13.3	3321	13	ADR07314	Full leng
44	829.4	13.2	3064	3	AAZ15575	Human pro
45	585.2	9.3	1011	2	AAQ94312	Tyrosine

ALIGNMENTS

RESULT 1

ADI32147
ID ADI32147 standard; CDNA; 6263 BP.

AC ADI32147;

DT 17-JUN-2004 (first entry)

DE Human CDNA #1473.

KW Human; gene; ss; immunological response; immunopathological condition;
KW Crohn's disease; asthma; ulcerative colitis; hypereosinophilia;
KW Irritable bowel syndrome; osteoarthritis; rheumatoid arthritis;
KW acute monocytic leukaemia; antiinflammatory; antiasthmatic; antiulcer;
KW osteopathic; antiarthritic; antirheumatic; cytostatic.

OS Homo sapiens.

PN US6607879-B1.

XX 19-AUG-2003.

XX 09-FEB-1998; 98US-00023655.

XX 09-FEB-1998; 98US-00023655.

PA (INCY-) INCYTE CORP.

PI Cocks BG, Stuart SG, Seilhamer JJ;

XX WPI; 2003-895307/82.

XX A composition comprising a plurality of cDNAs, useful for detecting
PT altered expression of genes in an immunological response or for
PT diagnosing and treating an immunopathology, e.g. Crohn's disease, asthma
PT or osteoarthritis.

XX Claim 1; SEQ ID NO 1473; 50pp; English.

XX The invention relates to a composition comprising a plurality of cDNAs
CC for detecting the altered expression of genes in an immunological
CC response. The invention also relates to a method of diagnosing or
CC monitoring the treatment of an immunopathological condition in a sample,

CC comprising obtaining nucleic acids from a sample, contacting the nucleic
CC acids of the sample with an array comprising the plurality of cDNAs under
CC conditions to form one or more hybridisation complexes, detecting the
CC hybridisation complexes and comparing the levels of the detected
CC hybridisation complexes with the level of hybridisation complexes
CC detected in a non-diseased sample, where an altered level of the detected
CC hybridisation complexes correlates with the presence of an
CC immunopathological condition. Also disclosed are an expression profile
CC comprising a microarray and a plurality of detectable complexes and a
CC method for identifying a plurality of polynucleotide probes. The cDNAs
CC are useful as hybridisable array elements in a microarray for monitoring
CC the expression of target polynucleotides. The microarray can be used in
CC the diagnosis of an immunopathology, such as Crohn's disease, asthma,
CC ulcerative colitis, hypersensitivity, irritable bowel syndrome,
CC osteoarthritis, rheumatoid arthritis or acute monocytic leukaemia, and in
CC identifying agents for the treatment of the diseases. The microarray may
CC also be used in drug discovery and development, toxicological and
CC carcinogenicity studies, forensics or pharmacogenomics. The composition
CC may also be used in purification of a subpopulation of mRNAs, cDNAs or
CC genomic fragments. This sequence represents a human cDNA of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification but was obtained in electronic format directly
CC from USPTO at seqdata.uspto.gov/sequence.html.

XX
SQ Sequence 6263 BP; 1870 A; 1449 C; 1445 G; 1499 T; 0 U; 0 Other;

Query Match 100.0%; Score 6263; DB 11; Length 6263;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTAACTCAAGGAGAGCTGTGGTGAACACCCGTTGGATCTAAAGAACAGCTCTGAAAG 60
DB 1 GCTAACTCAAGGAGAGCTGTGGTGAACACCCGTTGGATCTAAAGAACAGCTCTGAAAG 60
QY 61 TGTTCACAGCTGAAATTTTCAGATCGGACAGACTCGCTGCGGCTCCGGAGGCGAGTTCCTCA 120
DB 61 TGTTCACAGCTGAAATTTTCAGATCGGACAGACTCGCTGCGGCTCCGGAGGCGAGTTCCTCA 120
QY 121 AGCTGCTCGGACAGCTGCTGCGGAGCTGAGGATGCTGAGGATGCTGAGGATGCTGCTG 180
DB 121 AGCTGCTCGGACAGCTGCTGCGGAGCTGAGGATGCTGAGGATGCTGAGGATGCTGCTG 180
QY 181 CTGCTCTCTCACTTCTTCTCCGACAGGATGCTGAGACACCTCCAAAGTTTACACGAA 240
DB 181 CTGCTCTCTCACTTCTTCTCCGACAGGATGCTGAGACACCTCCAAAGTTTACACGAA 240
QY 241 CCCGTTGATCAGACAGGGGTCTCTGGGAGGTTGCTCTTTTCACTGCGCAAGCTACCGGA 300
DB 241 CCCGTTGATCAGACAGGGGTCTCTGGGAGGTTGCTCTTTTCACTGCGCAAGCTACCGGA 300
QY 301 GACCCAGACCTAAATTTGCTGGAAACAAAGGAAAGAAAGTCAAGTCAAGATTTT 360
DB 301 GACCCAGACCTAAATTTGCTGGAAACAAAGGAAAGAAAGTCAAGTCAAGATTTT 360
QY 361 GAGGTAAATAGAGTTTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 420
DB 361 GAGGTAAATAGAGTTTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 420
QY 421 CCGAGGGATGAGGCGCAATTTATGAATGTGTGCGCTCAAAATATGTGGGAGAAATAGTGA 480
DB 421 CCGAGGGATGAGGCGCAATTTATGAATGTGTGCGCTCAAAATATGTGGGAGAAATAGTGA 480
QY 481 TCCACACAGACTCAAGTTTTCGGGAGAGATCAAAATTCACGAGGCTTCCCTACCATTTGAC 540
DB 481 TCCACACAGACTCAAGTTTTCGGGAGAGATCAAAATTCACGAGGCTTCCCTACCATTTGAC 540
QY 541 ATGGGCCACAGTTTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 600
DB 541 ATGGGCCACAGTTTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 600
QY 601 AGTGGTAATCCGGATCCAGAAATCACTTGGTTTAAAGATTTCTTACTGTGGACACAAAGC 660
DB 601 AGTGGTAATCCGGATCCAGAAATCACTTGGTTTAAAGATTTCTTACTGTGGACACAAAGC 660

QY 661 AACCAACAATCGTCTGATTAAGCAGTTACGATCAGAAATCTATTGGTGTACACCAATAGA 720
DB 661 AACCAACAATCGTCTGATTAAGCAGTTACGATCAGAAATCTATTGGTGTACACCAATAGA 720
QY 721 GGAGCCCTTCAGATTGAGCAGAGTGAAGAGTCTTGACCAAGGAAATATGAGTGTGTGTC 780
DB 721 GGAGCCCTTCAGATTGAGCAGAGTGAAGAGTCTTGACCAAGGAAATATGAGTGTGTGTC 780
QY 781 ACCAACAAGCCGGGCACTCGCTATTCCGCTCTGCGCAATTTATATGTGAGAGAGCTGCGA 840
DB 781 ACCAACAAGCCGGGCACTCGCTATTCCGCTCTGCGCAATTTATATGTGAGAGAGCTGCGA 840
QY 841 GAAGTTGCGGCTGCCACCAAGATTTCTATCCACCCCAATATCATGAATCATGCGCA 900
DB 841 GAAGTTGCGGCTGCCACCAAGATTTCTATCCACCCCAATATCATGAATCATGCGCA 900
QY 901 GCGGGAAGCCGTTAATATCACTGTGTGCGGCTGGGTGACCAATGCCCTTATGTAAAGTGG 960
DB 901 GCGGGAAGCCGTTAATATCACTGTGTGCGGCTGGGTGACCAATGCCCTTATGTAAAGTGG 960
QY 961 ATGTTGGGGCAGAGAATCTGACACCTGAAAGATGATGCAATAGGAAGAAATGTGCTA 1020
DB 961 ATGTTGGGGCAGAGAATCTGACACCTGAAAGATGATGCAATAGGAAGAAATGTGCTA 1020
QY 1021 GAAGTTGCGGCTGCCACCAAGATTTCTATCCACCCCAATATCATGAATCATGCGCT 1080
DB 1021 GAAGTTGCGGCTGCCACCAAGATTTCTATCCACCCCAATATCATGAATCATGCGCT 1080
QY 1081 GTCAATTGAAGCAATAGCAGACAGTCACTGTCAAAGCCCTTACCACCAACCTCCAGGAACCTCT 1140
DB 1081 GTCAATTGAAGCAATAGCAGACAGTCACTGTCAAAGCCCTTACCACCAACCTCCAGGAACCTCT 1140
QY 1141 GTAGTGACCCGAGACAGCTACAAGCATCACATGACGTGGGACTCTGGGAAACCTCTGAG 1200
DB 1141 GTAGTGACCCGAGACAGCTACAAGCATCACATGACGTGGGACTCTGGGAAACCTCTGAG 1200
QY 1201 CCTGTTCTTATTACATAATTTAGCATTAACCTTAAACCTTCTGAGGAACCTTACAAAGAA 1260
DB 1201 CCTGTTCTTATTACATAATTTAGCATTAACCTTAAACCTTCTGAGGAACCTTACAAAGAA 1260
QY 1261 ATTGATGGGTGGGCGACCAACAGCTACAGTGTGCTGGACTTAAGTCTCTACTCGGATTAT 1320
DB 1261 ATTGATGGGTGGGCGACCAACAGCTACAGTGTGCTGGACTTAAGTCTCTACTCGGATTAT 1320
QY 1321 GAATTCAGGGTGTGCTGCTCAATTAACATTTGGGGGGGGGCTCTCCAGGAAACCTCTGCTA 1380
DB 1321 GAATTCAGGGTGTGCTGCTCAATTAACATTTGGGGGGGGGCTCTCCAGGAAACCTCTGCTA 1380
QY 1381 ACACAAACCTTCAGAGCAAGCACCATCCAGTGTCCCGAGGGATGTCCAGGCAAGATTTG 1440
DB 1381 ACACAAACCTTCAGAGCAAGCACCATCCAGTGTCCCGAGGGATGTCCAGGCAAGATTTG 1440
QY 1441 AGTTCGACCACTTTTGGTACAGTGGAGGAAACCTGAAAGGCAAAATGGAAGAGATCCAA 1500
DB 1441 AGTTCGACCACTTTTGGTACAGTGGAGGAAACCTGAAAGGCAAAATGGAAGAGATCCAA 1500
QY 1501 GGATATAGAGTTTATTAACAATGGATCCCACTCAACATGTCAACAATCTGGATGAACAC 1560
DB 1501 GGATATAGAGTTTATTAACAATGGATCCCACTCAACATGTCAACAATCTGGATGAACAC 1560
QY 1561 AATGTAGCTCAGACGCAAAATCACTACTATTGGCAACTTGTGCGCCCAAGAAACATATTCT 1620
DB 1561 AATGTAGCTCAGACGCAAAATCACTACTATTGGCAACTTGTGCGCCCAAGAAACATATTCT 1620
QY 1621 GTCAAAAGTCTGGCTTTTACCTCAATTTGGAGATGGTCCCTTTTCAAGTGACATCAAGTC 1680
DB 1621 GTCAAAAGTCTGGCTTTTACCTCAATTTGGAGATGGTCCCTTTTCAAGTGACATCAAGTC 1680
QY 1681 ATCACTCAGACAGAGTTACAGGGCAGCCTAAATTTCAAGCAGAAACCTGAGTCTGAA 1740
DB 1681 ATCACTCAGACAGAGTTACAGGGCAGCCTAAATTTCAAGCAGAAACCTGAGTCTGAA 1740

[illegible]

Db 3901 GTGGTGTCAATGGATCTGGATCCGACGCCAATCAGCGATGAAGAGGAGGCTTATCTGG 3960
 Qy 3961 GTTGTAGGTCTGTCTTGCAGTGGTCTTTATCATCTGCAATGTGTCATGTCTTCTTT 4020
 Db 3961 GTTGTAGGTCTGTCTTGCAGTGGTCTTTATCATCTGCAATGTGTCATGTCTTCTTT 4020
 Qy 4021 TATAAAGGAGAGGGGAGAGTCCGACTCTAGAAAAAGCAGCATACCGAAACAATAAGGAG 4080
 Db 4021 TATAAAGGAGAGGGGAGAGTCCGACTCTAGAAAAAGCAGCATACCGAAACAATAAGGAG 4080
 Qy 4081 ATCCCTTTACACCAACCCAAACAGACCCCTGTAGAACTGAGGGGCGCTTAACTTTCAAAACCG 4140
 Db 4081 ATCCCTTTACACCAACCCAAACAGACCCCTGTAGAACTGAGGGGCGCTTAACTTTCAAAACCG 4140
 Qy 4141 GGTATGGGTAGCCATCTCTCAATACCCATCTTGGAACTTGGCAACCAATTTGAAAGATTG 4200
 Db 4141 GGTATGGGTAGCCATCTCTCAATACCCATCTTGGAACTTGGCAACCAATTTGAAAGATTG 4200
 Qy 4201 AAAGCAAAATGAACTTTGAAGTTTCCAGGAATATCAGTCAATTTGACCCCTGGCCAGCAG 4260
 Db 4201 AAAGCAAAATGAACTTTGAAGTTTCCAGGAATATCAGTCAATTTGACCCCTGGCCAGCAG 4260
 Qy 4261 TTCACTTTGGGAACATTCAAACTTGGAAAGTAAACAAACCAAGAAATAGATACGCGAAATGTA 4320
 Db 4261 TTCACTTTGGGAACATTCAAACTTGGAAAGTAAACAAACCAAGAAATAGATACGCGAAATGTA 4320
 Qy 4321 ATCGCATATGATCAATCTCCGGGTTCTCTATCAGCTATAGAGGGATCCAGGAGTGCAC 4380
 Db 4321 ATCGCATATGATCAATCTCCGGGTTCTCTATCAGCTATAGAGGGATCCAGGAGTGCAC 4380
 Qy 4381 TATGTGAATGCCAATCATATAGATAGGTTATAGGAAGCAAAATGCTATATTTGCAACACAG 4440
 Db 4381 TATGTGAATGCCAATCATATAGATAGGTTATAGGAAGCAAAATGCTATATTTGCAACACAG 4440
 Qy 4441 GGATCTCTCCCGGAAACATTTGGGGACTTTTGGGAGATGATATGGGAACAACCGAGTGCC 4500
 Db 4441 GGATCTCTCCCGGAAACATTTGGGGACTTTTGGGAGATGATATGGGAACAACCGAGTGCC 4500
 Qy 4501 ACAGTTGTCTATGATGACAAAACCTAGAGAAAGATCAAGGTGAAGTGTGACCAGTATTGG 4560
 Db 4501 ACAGTTGTCTATGATGACAAAACCTAGAGAAAGATCAAGGTGAAGTGTGACCAGTATTGG 4560
 Qy 4561 CCTAGCAGAGCAGAGAAACCCAGGACTGCTTCAAGTAAAGCTGCTTGTATCTGTGGAG 4620
 Db 4561 CCTAGCAGAGCAGAGAAACCCAGGACTGCTTCAAGTAAAGCTGCTTGTATCTGTGGAG 4620
 Qy 4621 CTGGCCACATATTTGTTTGAACATTTGCACTTTTCAAGAAATGTTCAAGTGAAGAGAGA 4680
 Db 4621 CTGGCCACATATTTGTTTGAACATTTGCACTTTTCAAGAAATGTTCAAGTGAAGAGAGA 4680
 Qy 4681 GAAGTGAGACAAATTCAGTTTCAAGTGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 4740
 Db 4681 GAAGTGAGACAAATTCAGTTTCAAGTGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 4740
 Qy 4741 CCTTTTCTAGCTTTCTTACGTAGAGTCAAAACCTGTAACCCCTCCGATGCTGGTCCGATG 4800
 Db 4741 CCTTTTCTAGCTTTCTTACGTAGAGTCAAAACCTGTAACCCCTCCGATGCTGGTCCGATG 4800
 Qy 4801 GTTGTGCACTGAGTGGGGAGTTGGCCGAGCTGGTGTCTTCAATGATAGATGCCATG 4860
 Db 4801 GTTGTGCACTGAGTGGGGAGTTGGCCGAGCTGGTGTCTTCAATGATAGATGCCATG 4860
 Qy 4861 TTAGAAAGAAATTAAGCATGAAATACTGATAGATATTTATGGCCATGTAATTTATGAGA 4920
 Db 4861 TTAGAAAGAAATTAAGCATGAAATACTGATAGATATTTATGGCCATGTAATTTATGAGA 4920
 Qy 4921 GCCCAGAGAACTATATGGTTTCAAAACAGAAACCAATACATCTTTATCCATGATGCACTG 4980
 Db 4921 GCCCAGAGAACTATATGGTTTCAAAACAGAAACCAATACATCTTTATCCATGATGCACTG 4980
 Qy 4981 TTAGAAAGCAGTGTGTTGGAATAACCGAAGTGGCAGCTAGAAACTTGTATGCTCATATT 5040
 Db 4981 TTAGAAAGCAGTGTGTTGGAATAACCGAAGTGGCAGCTAGAAACTTGTATGCTCATATT 5040

Qy 5041 CAGAAGCTGACACAAATAGAAACGGGAGAGAAATGTACAGGAATGGAGCTCGAATTTAAG 5100
 Db 5041 CAGAAGCTGACACAAATAGAAACGGGAGAGAAATGTACAGGAATGGAGCTCGAATTTAAG 5100
 Qy 5101 CGTCTAGCCAGCTCAAAAGCTCACACCTCAAGGTTTATCAGTGCCCAATCTTCCATGTAAT 5160
 Db 5101 CGTCTAGCCAGCTCAAAAGCTCACACCTCAAGGTTTATCAGTGCCCAATCTTCCATGTAAT 5160
 Qy 5161 AAATTTCAAAATCGCCTTGTAAATATATGCCATATGAATCCCAAGGGTATGCTGCGAG 5220
 Db 5161 AAATTTCAAAATCGCCTTGTAAATATATGCCATATGAATCCCAAGGGTATGCTGCGAG 5220
 Qy 5221 CCTATCCGTGGAGTAGAAGATCTGATTAATCAATCAATGCCAGTTTATTTGATGGATACAGA 5280
 Db 5221 CCTATCCGTGGAGTAGAAGATCTGATTAATCAATCAATGCCAGTTTATTTGATGGATACAGA 5280
 Qy 5281 CAACAGAAAGCCTACATCGCTACCCAGGGGCGCTTGGCAGAGACCACTGAAGACTTCTGG 5340
 Db 5281 CAACAGAAAGCCTACATCGCTACCCAGGGGCGCTTGGCAGAGACCACTGAAGACTTCTGG 5340
 Qy 5341 CGGATGCTCTGGGAAACAATTCACCAATGCTCAATAGTTGTGATGCTCAACAGCTGCGTGAATG 5400
 Db 5341 CGGATGCTCTGGGAAACAATTCACCAATGCTCAATAGTTGTGATGCTCAACAGCTGCGTGAATG 5400
 Qy 5401 GGAGAGAGAAATGTCCCAATACTGCGCAGAGAAAGGCTCTGCAAGATACCAAGTACTT 5460
 Db 5401 GGAGAGAGAAATGTCCCAATACTGCGCAGAGAAAGGCTCTGCAAGATACCAAGTACTT 5460
 Qy 5461 GTTGTAGATCCCATGCTGAGTACCAATGCTCAAGTCAATCTTCAAGGAAATTCAGAGTC 5520
 Db 5461 GTTGTAGATCCCATGCTGAGTACCAATGCTCAAGTCAATCTTCAAGGAAATTCAGAGTC 5520
 Qy 5521 ACAGATGCTGAGGAGCGGCGAGTCCGAAACAGTAAAGGAGTTCCAGTTCACGTGCGCCA 5580
 Db 5521 ACAGATGCTGAGGAGCGGCGAGTCCGAAACAGTAAAGGAGTTCCAGTTCACGTGCGCCA 5580
 Qy 5581 GAGCAAGGAGTGCACAAAGTCCGAGAGAGGATTTTATGACCTTCAATCGGCAAGTCCATAA 5640
 Db 5581 GAGCAAGGAGTGCACAAAGTCCGAGAGAGGATTTTATGACCTTCAATCGGCAAGTCCATAA 5640
 Qy 5641 AAAAAAGAACAGTTTGGCCAAAGATGGAACCAATTTCACTGACGCGCGGGCTTGGGA 5700
 Db 5641 AAAAAAGAACAGTTTGGCCAAAGATGGAACCAATTTCACTGACGCGCGGGCTTGGGA 5700
 Qy 5701 AGAATCTGGAGTCTTCAATACGCTAAGCATTTGTTTGGAAAGAAATGAGATATGAAGGAGTT 5760
 Db 5701 AGAATCTGGAGTCTTCAATACGCTAAGCATTTGTTTGGAAAGAAATGAGATATGAAGGAGTT 5760
 Qy 5761 GTAGATATCTTCCAGACTGTCAAAAATGTTAAGAAACCAACGACGCTATGGTACAGACA 5820
 Db 5761 GTAGATATCTTCCAGACTGTCAAAAATGTTAAGAAACCAACGACGCTATGGTACAGACA 5820
 Qy 5821 GAGGATCAATATCAGTTTCTCTATCGTGCCGCACTAGAGTACCTGGGAGCTTTGACCAAC 5880
 Db 5821 GAGGATCAATATCAGTTTCTCTATCGTGCCGCACTAGAGTACCTGGGAGCTTTGACCAAC 5880
 Qy 5881 TATGCAACGTAAGAAAACCCCTGACCCATTTCTGGAATTTTACTACAGGCGCTTCAATATCCA 5940
 Db 5881 TATGCAACGTAAGAAAACCCCTGACCCATTTCTGGAATTTTACTACAGGCGCTTCAATATCCA 5940
 Qy 5941 TGGAGTCTCTTCTGAGCCATACAGGGCACTTGAGAGTCTCTTCTTAACTTCTAGCTAAACA 6000
 Db 5941 TGGAGTCTCTTCTGAGCCATACAGGGCACTTGAGGCACTTGAGAGTCTCTTCTTAACTTCTAGCTAAACA 6000
 Qy 6001 ACTACTTGTGGGACTATTACACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6060
 Db 6001 ACTACTTGTGGGACTATTACACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6060
 Qy 6061 AAGAAATCTTGTGACATCGCCCTTCCCACTACTGCTCATATAAATTTTGGGGGCA 6120
 Db 6061 AAGAAATCTTGTGACATCGCCCTTCCCACTACTGCTCATATAAATTTTGGGGGCA 6120

Qy	6121	AGGGAGGGAATGTTTTAAAAAAGAAGTCCTTGATTAGTTTGTAAAGATC	6180
Db	6121	AGGGAGGGAATGTTTTAAAAAAGAAGTCCTTGATTAGTTTGTAAAGATC	6180
Qy	6181	TGCTGACCTGTGCTTCATTTCTAACCTGTGTAAACTTTTTTTAACAAAATGTAATCATTCG	6240
Db	6181	TGCTGACCTGTGCTTCATTTCTAACCTGTGTAAACTTTTTTTAACAAAATGTAATCATTCG	6240
Qy	6241	ATAAAGTGAATTTTAAAAAAGTT	6263
Db	6241	ATAAAGTGAATTTTAAAAAAGTT	6263

RESULT 2

ADR40182
ID ADR40182 standard: cDNA: 6263 BP.

AC ADR40182:

18-NOV-2004 (first entry)

Human protein tyrosine phosphatase delta (R-PTP-delta) (965) cDNA.

haematological; cytostatic; erythroid; anaemia; erythrocytosis;
bone marrow; leukaemia; platelet; thrombocytopenia; thrombosis; B-cell;
T-cells; neutropenia; gene therapy; human; ss; gene;
protein tyrosine phosphatase delta; R-PTP-delta.

OS Homo sapiens.

Key	Location/Qualifiers
CDS	154. .5892

THE

EX

PN W020040/2242-AZ: XX

FD XX
26-AUG-2004.

XX
FF
2004-03-03

PK 03-FEB-2003; 2003US-044524IF.
PR 18-FEB-2003; 2003US-0448389P.

20-MAR-2003; 2003US-0438320F.
03-APR-2003; 2003US-0460279P.

EK Z8-AFK-2003; 2003US-0483924F.
 PR 13-MAY-2003; 2003US-0470052P.

FK Z6-AUG-2003; 2003US-0498106F;
PR 04-SEP-2003; 2003US-0500179F;

13-SEP-2003; 2003US-0510351P.
10-OCT-2003; 2003US-0510351P.
PR

FR 17-001-2003; 200305-0312380F:
XX

[illegible][illegible]

DR P-PSDB; ADR40183.

Identifying a compound capable

PT with the disorder under conditions suitable

XX
XX

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

CC capable of treating a haematological disorder which comprises combining
CC and invention relates to a novel method for identifying a compound
CC capable of treating a haematological disorder which comprises combining
CC compound to be tested with a specific polypeptide under conditions
CC suitable for binding of the test compound to the polypeptide. The method
CC of the invention has haematological and cycostatic applications and may

CC be useful for identifying compounds for treating a haematological
CC disorder associated with erythroid cells e.g. anaemia and erythrocytosis,
CC bone marrow e.g. leukaemia, platelets e.g. thrombocytopenia and
CC thrombosis or B-cells and T-cells e.g. neutropenia. The compounds
CC identified may be utilised during gene therapy procedures. The current
CC sequence is that of a human haematological disorder-related cDNA of the
CC invention.

Sequence 6263 BP: 1870 A: 1449 C: 1445 G: 1499 T: 0 U: 0 Other:

```
Query Match      100.0%; Score 6263; DB 13; Length 6263;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6263; Conservative 0; Mismatches 0; Indels 0;
```

Qy	1	GCTAACTCAAGGGAGACGTC	TGGTGAACACCCGTGGGATCTAAAGAACAAAGCTCTCAAG	60
Db	1	GCTAACTCAAGGGAGACGTC	TGGTGAACACCCGTGGGATCTAAAGAACAAAGCTCTCAAG	60
Qy	61	TGTTCCAGCTGAAATTTTCAGATCGGACAGAC	CTCGCTGCGGCTCCGGAGGCGAGTTCCTCA	120
Db	61	TGTTCCAGCTGAAATTTTCAGATCGGACAGAC	CTCGCTGCGGCTCCGGAGGCGAGTTCCTCA	120
Qy	121	AGTGTCTCGGCACCGTCTGTCGCAAGCTGCA	AGATGAGTGTGCTGCTG	180
Db	121	AGTGTCTCGGCACCGTCTGTCGCAAGCTGCA	AGTGTGCTGCTGCTG	180
Qy	181	CTGCTCCTCAGCTTTCTTCTCGCACGGATGCT	GAGACACCTTCCAAGGTTTACACGAACA	240
Db	181	CTGCTCCTCAGCTTTCTTCTCGCACGGATGCT	GAGACACCTTCCAAGGTTTACACGAACA	240
Qy	241	CCCGTTGATCAGACAGGGGTCTCTGGCGAGT	TGCTCTTTCATCTGCCAAGCTACGGGA	300
Db	241	CCCGTTGATCAGACAGGGGTCTCTGGCGAGT	TGCTCTTTCATCTGCCAAGCTACGGGA	300
Qy	301	GACCCAAAGCCTAAATTTCTTGGACAAAGAA	AGGAAAGTTCAGCAATCAGAGTTT	360
Db	301	GACCCAAAGCCTAAATTTCTTGGACAAAGAA	AGGAAAGTTCAGCAATCAGAGTTT	360
Qy	361	GAGGTAATAGAGTTTGACGATGGGTCTCGAT	CAGTTTCTCAGAAATACAACCTTTACGGACT	420
Db	361	GAGGTAATAGAGTTTGACGATGGGTCTCGAT	CAGTTTCTCAGAAATACAACCTTTACGGACT	420
Qy	421	CCGAGGGATGAGGCCATTTATGAATGTGTGG	CTCAATTAATGTGGGAGAAATAAGTGTA	480
Db	421	CCGAGGGATGAGGCCATTTATGAATGTGTGG	CTCAATTAATGTGGGAGAAATAAGTGTA	480
Qy	481	TCACACAGACTCACAGTTTTCGGGAAAGTCAA	ATTTCCAGAGGGCTTCCTTACCATTGAC	540
Db	481	TCACACAGACTCACAGTTTTCGGGAGAGTCAA	ATTTCCAGAGGGCTTCCTTACCATTGAC	540
Qy	541	ATGGGCCACACAGTTGAAAGTGTGAGCGTACT	CTCGACAGGCCACCAATGTTGTGCAGCC	600
Db	541	ATGGGCCACACAGTTGAAAGTGTGAGCGTACT	CTCGACAGGCCACCAATGTTGTGCAGCC	600
Qy	601	AGTGTGTATCCGGATCCGAAATCACTTGGTTT	AAAGATTTCTTACTCTGTGGACCAAGC	660
Db	601	AGTGTGTATCCGGATCCGAAATCACTTGGTTT	AAAGATTTCTTACTCTGTGGACCAAGC	660
Qy	661	AAACAACAATGGTTCGTATTTAAGCAGTTAC	GATCAGAAATCTATTGGTGGTACACCAATAAGA	720
Db	661	AAACAACAATGGTTCGTATTTAAGCAGTTAC	GATCAGAAATCTATTGGTGGTACACCAATAAGA	720
Qy	721	GGAGCCCTTCAGATTGAGCAGAGTGAAGTCTG	ACCAAGGAAAAATATGAGTGTGTGGC	780
Db	721	GGAGCCCTTCAGATTGAGCAGAGTGAAGTCTG	ACCAAGGAAAAATATGAGTGTGTGGC	780
Qy	781	ACCAACAGCGGGGCACTCGCTATTTCGGTCT	CTCGCAATTTATATGTTCAGAGAGCTGCGA	840
Db	781	ACCAACAGCGGGGCACTCGCTATTTCGGTCT	CTCGCAATTTATATGTTCAGAGAGCTGCGA	840
Qy	841	GAGTTGCGCGTGTCCCAACCAAGATTCTCTAT	CCCCACCCCAATCATGAAATCATGCCA	900
Db	841	GAGTTGCGCGTGTCCCAACCAAGATTCTCTAT	CCCCACCCCAATCATGAAATCATGCCA	900

Qy	901	GGCGAAGCGTTAATATACCTGTGTGCGCGTGGGGTCACCAATGCCCTTATGTAAAGTGG	960	1981	TCAGCTCCTCCTCAAGACATTAGTGTGCACAGCCCAAGTTCACACTAGTATTTTGGTAAGT	2040
Db	901	GGCGAAGCGTTAATATACCTGTGTGCGCGTGGGGTCACCAATGCCCTTATGTAAAGTGG	960	1981	TCAGCTCCTCCTCAAGACATTAGTGTGCACAGCCCAAGTTCACACTAGTATTTTGGTAAGT	2040
Qy	961	ATGTTGGGGGAGAAAGATCTGACACCTGAAGATGATATGCAATAGGAAGAAATGTGCTA	1020	2041	TGGCAACCTCCACAGTGGAAAAACAGAAATGGCATTATCACTGAATATCTCCATCAAGTAC	2100
Db	961	ATGTTGGGGGAGAAAGATCTGACACCTGAAGATGATATGCAATAGGAAGAAATGTGCTA	1020	2041	TGGCAACCTCCACAGTGGAAAAACAGAAATGGCATTATCACTGAATATCTCCATCAAGTAC	2100
Qy	1021	GAACCTGAATGATGTAAGACAGTCAAGCAATTTACACCTGTGTGCTATGTCAACACTGGGT	1080	2101	ACTGCAGTGGATGGGAAGATGACAGACCTCAGACAGATTTTGGGAATCTCTCGGCACCT	2160
Db	1021	GAACCTGAATGATGTAAGACAGTCAAGCAATTTACACCTGTGTGCTATGTCAACACTGGGT	1080	2101	ACTGCAGTGGATGGGAAGATGACAGACCTCAGACAGATTTTGGGAATCTCTCGGCACCT	2160
Qy	1081	GTCAATTGAAGCAATAGACAGATCACTGTCAAGCCTTACCCAAACCTCCAGGAACTCCT	1140	2161	ACCAAAATACCTTTTGGAAACAGCTGGAAAAATGACCTGAATACCGGATCACTGTGACAGCC	2220
Db	1081	GTCAATTGAAGCAATAGACAGATCACTGTCAAGCCTTACCCAAACCTCCAGGAACTCCT	1140	2161	ACCAAAATACCTTTTGGAAACAGCTGGAAAAATGACCTGAATACCGGATCACTGTGACAGCC	2220
Qy	1141	GTAGTGACCGAGAGACAGCTTACAGCATCACACTGACGTGGGACTCTGGGAACCCCTGAG	1200	2221	CATACAGATGTGCGGCCCTCGCCCTGAGAGCTTGTGCGTGTGATTCGAACCAATGAAGAT	2280
Db	1141	GTAGTGACCGAGAGACAGCTTACAGCATCACACTGACGTGGGACTCTGGGAACCCCTGAG	1200	2221	CATACAGATGTGCGGCCCTCGCCCTGAGAGCTTGTGCGTGTGATTCGAACCAATGAAGAT	2280
Qy	1201	CCTGTTCTTTATTAATAATTCAGCATTAACCTTAACCTTAACCTTAACCTTAACCTTAAC	1260	2281	GTTCCTAGTGGTCTCTCTCGCAAGTGCAGGTAGAGCTGTCAACTCAACATCTGTTAAA	2340
Db	1201	CCTGTTCTTTATTAATAATTCAGCATTAACCTTAACCTTAACCTTAACCTTAACCTTAAC	1260	2281	GTTCCTAGTGGTCTCTCTCGCAAGTGCAGGTAGAGCTGTCAACTCAACATCTGTTAAA	2340
Qy	1261	ATTGATGGGGTGGCGACACACGCTACAGTGTGCTGCTGACTTAAGTCCCTACTCGGATTAT	1320	2341	GTCTCATGCGCTCACCCGTGCCCCAATAAACAGCATGGCCAGATAAGAGGATATCAGGTG	2400
Db	1261	ATTGATGGGGTGGCGACACACGCTACAGTGTGCTGCTGACTTAAGTCCCTACTCGGATTAT	1320	2341	GTCTCATGCGCTCACCCGTGCCCCAATAAACAGCATGGCCAGATAAGAGGATATCAGGTG	2400
Qy	1321	GAAATTCAGGGTGTGTGCTCAATAAATTTGGGCGGGGGCTCCACAGCGAACCTGTGCTA	1380	2401	CATTATGTGAGGATGGAATAATGTGAGCCCAAGGGCCAGCCCATGTCTGAAAGATGTGATG	2460
Db	1321	GAAATTCAGGGTGTGTGCTCAATAAATTTGGGCGGGGGCTCCACAGCGAACCTGTGCTA	1380	2401	CATTATGTGAGGATGGAATAATGTGAGCCCAAGGGCCAGCCCATGTCTGAAAGATGTGATG	2460
Qy	1381	ACACAAACCTCAGAGCAAGCAACCTACAGTGTGCTGCGGAGATGTCCAGGCAAGATGTTG	1440	2461	CTGGCTGATGCAAGTGGGAATTTGATGATCTACTGAAACATGACATGATCTTCTGGG	2520
Db	1381	ACACAAACCTCAGAGCAAGCAACCTACAGTGTGCTGCGGAGATGTCCAGGCAAGATGTTG	1440	2461	CTGGCTGATGCAAGTGGGAATTTGATGATCTACTGAAACATGACATGATCTTCTGGG	2520
Qy	1441	AGTTGACACCAATTTGGTACAGTGGAGAACTGAGAGCCAAATGAGCAGATCCAA	1500	2521	CTCAGCCTGAAACTTCTACTCCCTCACCGTCAACCGTCAACCGTCAACCGTCAACCGTCAAC	2580
Db	1441	AGTTGACACCAATTTGGTACAGTGGAGAACTGAGAGCCAAATGAGCAGATCCAA	1500	2521	CTCAGCCTGAAACTTCTACTCCCTCACCGTCAACCGTCAACCGTCAACCGTCAACCGTCAAC	2580
Qy	1501	GGATATAGAGTTTATATACAAATGGATCCCACTCAACATGTCAACAACTGGATGAACAC	1560	2581	GCTCGCAGCAAGCCCAAACTGGTGTCCACCACTCGGGCAGTTCAGGGGAAACCTCGGCTT	2640
Db	1501	GGATATAGAGTTTATATACAAATGGATCCCACTCAACATGTCAACAACTGGATGAACAC	1560	2581	GCTCGCAGCAAGCCCAAACTGGTGTCCACCACTCGGGCAGTTCAGGGGAAACCTCGGCTT	2640
Qy	1561	AATGTAGCTGACAGCAAAATCACTATTTGGGAACTTTAGTGTGCGCCAGAAACATATTTCT	1620	2641	GTGATTTAACCACTCAGATGAAATCTGCTCTTATTTCAGTGGGACCCCTCCGGTGGACACA	2700
Db	1561	AATGTAGCTGACAGCAAAATCACTATTTGGGAACTTTAGTGTGCGCCAGAAACATATTTCT	1620	2641	GTGATTTAACCACTCAGATGAAATCTGCTCTTATTTCAGTGGGACCCCTCCGGTGGACACA	2700
Qy	1621	GTCAAGTCTGCTGCTTTTACCTCAATTTGGAGATGGTCCCTTTCAAGTGACATACAAAGTC	1680	2701	TTTGGACCTCTTCAGGGCTACCGTCTAAAAATTTGGCCGCAAGGATATGAGGCCACTTTACT	2760
Db	1621	GTCAAGTCTGCTGCTTTTACCTCAATTTGGAGATGGTCCCTTTCAAGTGACATACAAAGTC	1680	2701	TTTGGACCTCTTCAGGGCTACCGTCTAAAAATTTGGCCGCAAGGATATGAGGCCACTTTACT	2760
Qy	1681	ATCACTCAGACAGGATACAGGGCAGCCACTAAATTTCAAGCAGAACCTGAGTCTGAA	1740	2761	ACTCTTGAGTCTCTGAAAAAGAGATCACTTTTACAGCTACAGACATCCACAGGGAGCA	2820
Db	1681	ATCACTCAGACAGGATACAGGGCAGCCACTAAATTTCAAGCAGAACCTGAGTCTGAA	1740	2761	ACTCTTGAGTCTCTGAAAAAGAGATCACTTTTACAGCTACAGACATCCACAGGGAGCA	2820
Qy	1741	ACAAAGTATTTTGTCTCTTTGACACCTTCAAGTGTGAGATGCTTCAAGTGTGAGTGTGAG	1800	2821	TCATAGCTCTTCAGGCTCTCAGCCAGAAACAAAGTGGCTTTTGGGGAGGAGATGTTGAAG	2880
Db	1741	ACAAAGTATTTTGTCTCTTTGACACCTTCAAGTGTGAGATGCTTCAAGTGTGAGTGTGAG	1800	2821	TCATAGCTCTTCAGGCTCTCAGCCAGAAACAAAGTGGCTTTTGGGGAGGAGATGTTGAAG	2880
Qy	1801	GTCTACAAAGATGGGAGCATGGAGGAGCAACGAAATTTACATTTGAGCCAGGAGCATCA	1860	2881	GAGATTTTCCATTCAGAAAGAGTACCAACTGGATTCCTCAAAACCTTCACTCAGAAAGGC	2940
Db	1801	GTCTACAAAGATGGGAGCATGGAGGAGCAACGAAATTTACATTTGAGCCAGGAGCATCA	1860	2881	GAGATTTTCCATTCAGAAAGAGTACCAACTGGATTCCTCAAAACCTTCACTCAGAAAGGC	2940
Qy	1861	TATAGGCTGCAAGGACTGAAACCAACAGCTTATATTTTCCGCTGTGGCTGACGCTCC	1920	2941	ACCACTTCAACCTCCGTCAGTTATCTTGGCAACCACTGTCTCTGCGCAGAGAGAAATGGC	3000
Db	1861	TATAGGCTGCAAGGACTGAAACCAACAGCTTATATTTTCCGCTGTGGCTGACGCTCC	1920	2941	ACCACTTCAACCTCCGTCAGTTATCTTGGCAACCACTGTCTCTGCGCAGAGAGAAATGGC	3000
Qy	1921	CCTCAAGGCTGGGTGCTTCTTACAGCAAAATATACAGCTAGAACCAATGAGTCAAGCCG	1980	3001	ATTATACCAAGTATACCTTCTTTTATAGGATATCAACATCCCCCTTCTCCCGATGGAG	3060
Db	1921	CCTCAAGGCTGGGTGCTTCTTACAGCAAAATATACAGCTAGAACCAATGAGTCAAGCCG	1980	3001	ATTATACCAAGTATACCTTCTTTTATAGGATATCAACATCCCCCTTCTCCCGATGGAG	3060
				3061	CAGCTTATGTTTCAGCTGACCAACACTATGACACTCACTGGCTTTAAAAACCAAGATACCA	3120

Db 3061 |||||CAGCTTATGTTCCAGCTGACACCACTATGACACTCACTGGCTTAAACCCAGATACCCACA||| 3120
Qy 3121 TACGATGTAAAGTACGTGCTCATACGAGCAAAAGGCGCGGCCCATATAGTCCCAAGTGTG 3180
Db 3121 TACGATGTAAAGTACGTGCTCATACGAGCAAAAGGCGCGGCCCATATAGTCCCAAGTGTG 3180
Qy 3181 CAGTTTCAGGACATGCGCTGTGGATCMAGTGTGTCGAAAATTTTCATGTCAAAGCAGTA 3240
Db 3181 CAGTTTCAGGACATGCGCTGTGGATCAAGTGTGTCGAAAATTTTCATGTCAAAGCAGTA 3240
Qy 3241 ATGAAGACTTCCGTTGCTGCTTTGGAGATTTCCAGAGAAATTAATCTCCGCCATGCGCT 3300
Db 3241 ATGAAGACTTCCGTTGCTGCTTTGGAGATTTCCAGAGAAATTAATCTCCGCCATGCGCT 3300
Qy 3301 TTCAAAATTTCTTATGATGATGGGAAAATGGTAGAAGATGGATGGCGGAGCCACACAG 3360
Db 3301 TTCAAAATTTCTTATGATGATGGGAAAATGGTAGAAGATGGATGGCGGAGCCACACAG 3360
Qy 3361 AAGTTAATGTCACCTGAAGCTGAGAAATCATATTCATTTGTGCTGACAAAATGCTGGA 3420
Db 3361 AAGTTAATGTCACCTGAAGCTGAGAAATCATATTCATTTGTGCTGACAAAATGCTGGA 3420
Qy 3421 AACAGTGTGCTGSGCTCAGCAGAGGCTCAGGCAAGAGCTGCACCAAGATGATTTACGT 3480
Db 3421 AACAGTGTGCTGSGCTCAGCAGAGGCTCAGGCAAGAGCTGCACCAAGATGATTTACGT 3480
Qy 3481 ACCAAGCTGCTTCATTTGGGAAGACCAATTTGGATGGCATGATTTACTGTGCAACTGCGCT 3540
Db 3481 ACCAAGCTGCTTCATTTGGGAAGACCAATTTGGATGGCATGATTTACTGTGCAACTGCGCT 3540
Qy 3541 GAAGTACCTGCAAAATGAGAAATATAAAGGTTTACTACATAATAATTTGTGCTTTGAAAGAA 3600
Db 3541 GAAGTACCTGCAAAATGAGAAATATAAAGGTTTACTACATAATAATTTGTGCTTTGAAAGAA 3600
Qy 3601 TCTCGCGGAAATTTATCAAGCCATGGGAGAGTCCAGATGAATGGAATAGATGAGCTG 3660
Db 3601 TCTCGCGGAAATTTATCAAGCCATGGGAGAGTCCAGATGAATGGAATAGATGAGCTG 3660
Qy 3661 CTTAAGGAGATATCTAGGAAGCGAGAGCATCCGTTATGGGAGAGAGTTGAATTAAG 3720
Db 3661 CTTAAGGAGATATCTAGGAAGCGAGAGCATCCGTTATGGGAGAGAGTTGAATTAAG 3720
Qy 3721 CCATATATGCGGCTCACTTTGATGCTCTCCCACTGAGTTTCACTCCGGGGATGACAAG 3780
Db 3721 CCATATATGCGGCTCACTTTGATGCTCTCCCACTGAGTTTCACTCCGGGGATGACAAG 3780
Qy 3781 CATTATGTTGATTTACAAACAAGCACTCCAAAGTGGTCAAGATATGCTCTTTGTTG 3840
Db 3781 CATTATGTTGATTTACAAACAAGCACTCCAAAGTGGTCAAGATATGCTCTTTGTTG 3840
Qy 3841 TTAGCAGTAAATGGAAACATGCAAGTCTAAGATGTATGCAACAGCCCTTACTCCGACCCC 3900
Db 3841 TTAGCAGTAAATGGAAACATGCAAGTCTAAGATGTATGCAACAGCCCTTACTCCGACCCC 3900
Qy 3901 GTGTGTCAATGATTTGATTCGAGCCCAATCAAGATGAAGAGAGGCTTGTATCTGG 3960
Db 3901 GTGGTGTCAATGATTCGATTCGAGCCCAATCAAGATGAAGAGAGGCTTGTATCTGG 3960
Qy 3961 GTTGTAGTCTGTCCTGTCAGTGGTCTTTTATCATCTGCAATGTGTCATTCTTCTTCTT 4020
Db 3961 GTTGTAGTCTGTCCTGTCAGTGGTCTTTTATCATCTGCAATGTGTCATTCTTCTTCTT 4020
Qy 4021 TATAAAGGAGAGAGGCGAGTCCGACTCTAGAAAAAGCAGCATACCGAAACAATAAGGAG 4080
Db 4021 TATAAAGGAGAGAGGCGAGTCCGACTCTAGAAAAAGCAGCATACCGAAACAATAAGGAG 4080
Qy 4081 ATCCCTTTCACACCAACCAAGAGCCCTGTGTAAGCTGAGGCGCCCTTAACTTTCAAAACCG 4140
Db 4081 ATCCCTTTCACACCAACCAAGAGCCCTGTGTAAGCTGAGGCGCCCTTAACTTTCAAAACCG 4140
Qy 4141 GGTATGCTAGCCATCTCTCAATACCCATCTTGGAACTTGCAGACCAATTTGAAGATTG 4200

Db 4141 GGTATGCTAGCCATCTCTCAATACCCATCTTGGAACTTGCAGACCAATTTGAAGATTG 4200
Qy 4201 AAAGCAAAATGACAACTTTGAAAGTTTTCCAGGAATATGAGTCAATTTGACCTCGGCAGCAG 4260
Db 4201 AAAGCAAAATGACAACTTTGAAAGTTTTCCAGGAATATGAGTCAATTTGACCTCGGCAGCAG 4260
Qy 4261 TTCACTTTGGGAACATTTCAAACCTTGGGAATTAACCAACCAAGAAATAGATACCGGAATGTA 4320
Db 4261 TTCACTTTGGGAACATTTCAAACCTTGGGAATTAACCAACCAAGAAATAGATACCGGAATGTA 4320
Qy 4321 ATCCCATATGATCATTTCCCGGTTCTCTATCAGCTATAGAGGATCCCAAGGAAGTCAAC 4380
Db 4321 ATCCCATATGATCATTTCCCGGTTCTCTATCAGCTATAGAGGATCCCAAGGAAGTCAAC 4380
Qy 4381 TATGTGAATGCCAACTACATAGATGGGTATAGGAAGCAAAATGCTATTTATGCAACACAG 4440
Db 4381 TATGTGAATGCCAACTACATAGATGGGTATAGGAAGCAAAATGCTATTTATGCAACACAG 4440
Qy 4441 GGATCTCTCCCGGAAAACATTTTGGGGACTTTTGGGAGATGATATGGGAACAACGGAGTGCC 4500
Db 4441 GGATCTCTCCCGGAAAACATTTTGGGGACTTTTGGGAGATGATATGGGAACAACGGAGTGCC 4500
Qy 4501 ACAGTTCTCATGATGACAAAACCTAGAAAAGATCAAGGGTGAAGTGTGACCAAGTATTGG 4560
Db 4501 ACAGTTCTCATGATGACAAAACCTAGAAAAGATCAAGGGTGAAGTGTGACCAAGTATTGG 4560
Qy 4561 CCTAGCAGAGCCACAGAAACCCACGGAATCGTTCAAGTAAAGTGGTTCAGTGAAGAGAGA 4620
Db 4561 CCTAGCAGAGCCACAGAAACCCACGGAATCGTTCAAGTAAAGTGGTTCAGTGAAGAGAGA 4620
Qy 4621 CTGGCCACATATTTGTTGTTGCAACATTTTGCACTTTTACAAGAAATGGTTCAGTGAAGAGAGA 4680
Db 4621 CTGGCCACATATTTGTTGTTGCAACATTTTGCACTTTTACAAGAAATGGTTCAGTGAAGAGAGA 4680
Qy 4681 GAAGTGAGACAAATTTCCAGTTTCAAGTTCAGGCTGCTGATCATGTTTCCAGAACACCCCTACA 4740
Db 4681 GAAGTGAGACAAATTTCCAGTTTCAAGTTCAGGCTGCTGATCATGTTTCCAGAACACCCCTACA 4740
Qy 4741 CCTTTTCTAGCTTTCTTACGTAGAGTCAAAACCTGTAACCTCCGATGCTGGTCCGATG 4800
Db 4741 CCTTTTCTAGCTTTCTTACGTAGAGTCAAAACCTGTAACCTCCGATGCTGGTCCGATG 4800
Qy 4801 GTTGTGACATGCAAGTGGGAGTTGGCGGAGTGGTTCATCGTCATAGATGCCATG 4860
Db 4801 GTTGTGACATGCAAGTGGGAGTTGGCGGAGTGGTTCATCGTCATAGATGCCATG 4860
Qy 4861 TTAGAAAGAAATAAGCATGAAAACCTGATAGATTTATGGCCATGTAACCTTTAATGAGA 4920
Db 4861 TTAGAAAGAAATAAGCATGAAAACCTGATAGATTTATGGCCATGTAACCTTTAATGAGA 4920
Qy 4921 GCCCAGAGGAACATATATGTTTCAAAACAGAGAACCAATACATCTTTATCCATGATGCACTG 4980
Db 4921 GCCCAGAGGAACATATATGTTTCAAAACAGAGAACCAATACATCTTTATCCATGATGCACTG 4980
Qy 4981 TTAGAACAGTACTTGTGGAATAACCGAAGTGGCAGCTAGAAAACCTTGTATGCCATACATT 5040
Db 4981 TTAGAACAGTACTTGTGGAATAACCGAAGTGGCAGCTAGAAAACCTTGTATGCCATACATT 5040
Qy 5041 CAGAACGTACACAAATAGAAAACCGGAGAGAAATGTCACAGAAATGGAGCTCGAATTTAAG 5100
Db 5041 CAGAACGTACACAAATAGAAAACCGGAGAGAAATGTCACAGAAATGGAGCTCGAATTTAAG 5100
Qy 5101 CGTCTAGCCAGCTCAAAAGCTCAACCTCAAGGTTTATCAGTGCCTCAATCTTCCATGTAAT 5160
Db 5101 CGTCTAGCCAGCTCAAAAGCTCAACCTCAAGGTTTATCAGTGCCTCAATCTTCCATGTAAT 5160
Qy 5161 AAATTCAAAATCGCCTTGTAAATATATGCAATGCAATCCCAAGGGTATGCTCGCAG 5220
Db 5161 AAATTCAAAATCGCCTTGTAAATATATGCAATGCAATCCCAAGGGTATGCTCGCAG 5220
Qy 5221 CCTATCCGTGAGTGAAGGATCTGATTACATCAATGCCAGTTTATTTGATGGATACAGA 5280
Db 5221 CCTATCCGTGAGTGAAGGATCTGATTACATCAATGCCAGTTTATTTGATGGATACAGA 5280

QY 5281 CAACAGAAAGCTACATCGCTACCCAGGGGCCCTTGGCAGAGACCACTGAGACTTCTGG 5340
 Db 5281 CAACAGAAAGCTACATCGCTACCCAGGGGCCCTTGGCAGAGACCACTGAGACTTCTGG 5340
 QY 5341 CGGATGCTCTGGGAACAACATTCACCATAGTGTGATGCTCACCAGAGCTGCGTGAATG 5400
 Db 5341 CGGATGCTCTGGGAACAACATTCACCATAGTGTGATGCTCACCAGAGCTGCGTGAATG 5400
 QY 5401 GGCAGAGAAATGTACCAATATCTGGCCAGAGAAAGGTCTGCAAGATACCACTACTTT 5460
 Db 5401 GGCAGAGAAATGTACCAATATCTGGCCAGAGAAAGGTCTGCAAGATACCACTACTTT 5460
 QY 5461 GTTCTAGATCCATGGCTGAGTACCAATGCGCAGTATATCTTAAGGAAATCAAGGTC 5520
 Db 5461 GTTCTAGATCCATGGCTGAGTACCAATGCGCAGTATATCTTAAGGAAATCAAGGTC 5520
 QY 5521 ACAGATGCCAGGAGCGGCAGTCCCGAACAGATAGAGGAGTTCACAGTTCACCTGAGTGGCCA 5580
 Db 5521 ACAGATGCCAGGAGCGGCAGTCCCGAACAGATAGAGGAGTTCACAGTTCACCTGAGTGGCCA 5580
 QY 5581 GAGCAAGAGTGCACAAAGTCCGGAAGAGATTTATTTGACTTCATCGGCCAAGTCCATAAA 5640
 Db 5581 GAGCAAGAGTGCACAAAGTCCGGAAGAGATTTATTTGACTTCATCGGCCAAGTCCATAAA 5640
 QY 5641 ACAGAGAAAGTGTGGCCAGATGGACCCATTTTCACTTCAGTCCAGCGGGGGTGGGA 5700
 Db 5641 ACAGAGAAAGTGTGGCCAGATGGACCCATTTTCACTTCAGTCCAGCGGGGGTGGGA 5700
 QY 5701 AGAAGTGGAGTCTTCATACAGCTAAGCATTTGTTTGGAAAGATGAGATATGAAGGAGTT 5760
 Db 5701 AGAAGTGGAGTCTTCATACAGCTAAGCATTTGTTTGGAAAGATGAGATATGAAGGAGTT 5760
 QY 5761 GTAGATATCTTCAGACTGTCAAAATGTTAAGAACACCAAGCAGCTATGGTACAGACA 5820
 Db 5761 GTAGATATCTTCAGACTGTCAAAATGTTAAGAACACCAAGCAGCTATGGTACAGACA 5820
 QY 5821 GAGGATCAATATCAGTTTTCTATCGTCCGACCTAGAGTACCTGGGCGGCTTTGACAC 5880
 Db 5821 GAGGATCAATATCAGTTTTCTATCGTCCGACCTAGAGTACCTGGGCGGCTTTGACAC 5880
 QY 5881 TATGCAAGTGAAGAACCCCTGACCCATTTCTGGATTTTACTACAGGCCCTTCAATATCCA 5940
 Db 5881 TATGCAAGTGAAGAACCCCTGACCCATTTCTGGATTTTACTACAGGCCCTTCAATATCCA 5940
 QY 5941 TGGAGTCTCTTCTGAGCCATACAGGCACTTGGAGAGTCTTCTTAACCTTCTAGCTAACA 6000
 Db 5941 TGGAGTCTCTTCTGAGCCATACAGGCACTTGGAGAGTCTTCTTAACCTTCTAGCTAACA 6000
 QY 6001 ACTACTTGTGGAGTATTACACACAAACAAATTTAAACAAATTTATTCAGGTGGACC 6060
 Db 6001 ACTACTTGTGGAGTATTACACACAAACAAATTTAAACAAATTTATTCAGGTGGACC 6060
 QY 6061 AAGAAATCTTTGACATCGCCCTTCCCAACCATCTGCTCATAATAACATTTAGGGGCCA 6120
 Db 6061 AAGAAATCTTTGACATCGCCCTTCCCAACCATCTGCTCATAATAACATTTAGGGGCCA 6120
 QY 6121 AGGGAGGGAATGTTTAAAGAAAGTCTTGAATTTAGTATTTTGAATTTGAAGATAC 6180
 Db 6121 AGGGAGGGAATGTTTAAAGAAAGTCTTGAATTTAGTATTTTGAATTTGAAGATAC 6180
 QY 6181 TGCTGACCTGTGCTTCAATTTCTTAAGTGTAAACCTTTTTTAAACAAATGTATCATTCG 6240
 Db 6181 TGCTGACCTGTGCTTCAATTTCTTAAGTGTAAACCTTTTTTAAACAAATGTATCATTCG 6240
 QY 6241 ATAAAGTGAATTTTAAAGAAAGTT 6263
 Db 6241 ATAAAGTGAATTTTAAAGAAAGTT 6263

RESULT 3
 ADD18743
 ID ADD18743 standard; DNA; 6363 BP.

XX ADD18743;
 AC AC
 XX DT 15-JAN-2004 (first entry)
 XX DE Human disease related protein DNA sequence SeqID174.
 XX KW human; disease state; cytostatic; antiinflammatory; ophthalmological;
 KW antiarteriosclerotic; vulnerary; gene therapy;
 KW hypoxia-regulated condition; tumorigenesis; angiogenesis; apoptosis;
 KW inflammation; erythropoiesis; glycolysis; gluconeogenesis;
 KW glucose transportation; catecholamine synthesis; iron transport;
 KW nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;
 KW retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;
 KW inflammatory condition; wound healing; gene; ds.
 XX OS Homo sapiens.
 OS WO2003018621-A2.
 PN 06-MAR-2003.
 XX PD 23-AUG-2002; 2002WO-GB003892.
 XX PF 23-AUG-2001; 2001GB-00020558.
 XX PR 05-OCT-2001; 2001GB-00024037.
 XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX PI Kingman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;
 XX WPI: 2003-290046/28.
 XX DR P-PSDB; ADD18742.
 XX PT New substantially purified polypeptide, useful for diagnosing or treating
 PT a hypoxia-regulated condition, such as cancer, ischemia, reperfusion
 PT injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or
 PT wound healing.
 XX PS Claim 27; SEQ ID NO 174; 424pp; English.
 XX CC This invention relates to novel human genes and gene product which are
 CC implicated in certain disease states. Compounds which modulate the
 CC proteins of the invention may have cytostatic, antiinflammatory,
 CC ophthalmological, antiarteriosclerotic or vulnerary activities. The
 CC sequences of the invention may be useful for gene therapy. The invention
 CC may be useful for diagnosing or treating a hypoxia-regulated condition,
 CC such as tumorigenesis, angiogenesis, apoptosis, inflammation,
 CC erythropoiesis, or the biological response to hypoxia conditions
 CC including processes such as glycolysis, gluconeogenesis, glucose
 CC transportation, catecholamine synthesis, iron transport or nitric oxide
 CC synthesis. The disease includes cancer, ischaemic conditions, reperfusion
 CC injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,
 CC inflammatory conditions or wound healing. The present sequence is that of
 CC a disease related protein encoding DNA sequence of the invention.
 XX SQ Sequence 6363 BP; 1301 A; 2073 C; 1921 G; 1068 T; 0 U; 0 Other;
 Query Match 39.5%; Score 2473; DB 10; Length 6363;
 Best Local Similarity 65.5%; Pred. No. 0;
 Matches 3821; Conservative 0; Mismatches 1915; Indels 99; Gaps 10;
 QY 149 GCAGGATGGTGCAGTACGAGCGCTGCTGCTGCTCCTCCTCCTCTTCTTCCTCCGACGG 208
 Db 131 GCATGGTGTCTGTGGTGTGGTCCCATGGGCTCTCTTGTGCTCTGCTGTTGGAGGCTGTG 190
 QY 209 ATGCTGAGACACCTCCAAAGGTTTACAGAACACCCGTTGATCAGACAGGGGTCTCTGGCG 268
 Db 191 CAGCAGAAGAGCCGCCCCAGGTTTATCAAGAACCCAGGACCCAGATCGGCTGTGCGGGG 250
 QY 269 GAGTGGCTCTTTTCATCTGCAAGCTACGGAGACCCCAAGACCTAAATTTGCTTGGGACA 328
 Db 251 GTGTGGCTCTTTTCGTGTGTGTCAGGCCCGGTGACCCCAAGCCACGAGTACCTTGGGACA 310

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 5, 2005, 20:04:47 ; Search time 25780 Seconds
(without alignments)
11771.723 Million cell updates/sec

Title: US-10-772-636-63
Perfect score: 6263
Sequence: 1 gctaactcaaggagagcgc.....aagtgaattttaaaaaagtt 6263

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6263	100.0	6263	6 AR380928	AR380928 Sequence
2	6263	100.0	6263	9 HUMPTPD	L38929 Homo sapien
3	5657.6	90.3	5754	6 CQ722829	CQ722829 Sequence
4	4584	73.2	4584	9 HSPPTD	X54133 Human HPTP
5	3856	61.6	6562	5 AF197944	AF197944 Xenopus 1
6	2473	39.5	6363	6 AR380334	AR380334 Sequence
7	2473	39.5	6363	9 HSU35234	U35234 Human prote
8	2418.4	38.6	4031	10 AF326559	AF326559 Mus muscu
9	2399.6	38.3	6000	9 HSU40317	U40317 Human prote
10	2398	38.3	4008	10 AF326560	AF326560 Mus muscu
11	2398	38.3	6000	6 AR031690	AR031690 Sequence
12	2398	38.3	6000	6 I61404	I61404 Sequence 6
13	2371	37.9	5723	10 AF300943	AF300943 Mus muscu
14	2365.4	37.8	4555	6 E09890	E09890 Mouse cdna
15	2357.4	37.6	7702	6 CQ871861	CQ871861 Sequence
16	2357.4	37.6	7702	6 AR380791	AR380791 Sequence
17	2357.4	37.6	7702	6 AX658135	AX658135 Sequence
18	2357.4	37.6	7702	9 HSLARR	Y00815 Human mRNA
19	2357.4	37.6	7945	6 CQ496043	CQ496043 Sequence

20	2346.6	37.5	6545	10 RATLARA	L11586 Rat leukocy
21	2321	37.1	5978	10 MMPTPN03	X82288 M.musculus
22	2320.8	37.1	4555	10 MUSMPTPA	D13903 Mus musculu
23	2315.8	37.0	6412	9 BC048768	BC048768 Homo sapi
24	2307.8	36.8	6734	6 AX305472	AX305472 Sequence
25	2307.8	36.8	6734	10 MUSPTPT9	D28530 Mouse mRNA
26	2173.8	34.7	6469	10 RATLARPTPA	L11587 Rat leukocy
27	2137.2	34.1	5634	6 CQ721109	CQ721109 Sequence
28	2118.8	33.8	2406	10 BC025145	BC025145 Mus muscu
29	2104.8	33.6	5424	6 CQ722132	CQ722132 Sequence
30	2025.2	32.3	7297	5 AF197945	AF197945 Xenopus 1
31	1928.6	30.8	5435	10 BC057166	BC057166 Mus muscu
32	1555.8	24.8	5953	9 HSM905669	BS337361 Homo sapi
33	1517.6	24.2	4836	6 CQ842367	-CQ842367 Sequence
34	1517.6	24.2	4836	9 AK125325	AK125325 Homo sapi
35	1393.2	22.2	4613	5 CHKCRYP	AR32780 Gallus gall
36	1345.8	21.5	4078	6 AR270569	AR270569 Sequence
37	1338.8	21.4	4773	9 HSU41725	U41725 Human prote
38	1338.8	21.4	5490	6 BD183305	BD183305 Novel gen
39	1301.2	20.8	5612	10 BC083188	BC083188 Mus muscu
40	1300	20.8	5690	6 AR060681	AR060681 Sequence
41	1300	20.8	5690	6 AR064169	AR064169 Sequence
42	1298.4	20.7	5383	10 RATLARPTPB	L12329 Rat leukocy
43	1296.8	20.7	4724	10 RATTYRPHOS	L19933 Rattus norv
44	1280.8	20.5	5588	10 BC052462	BC052462 Mus muscu
45	1245.2	19.9	5412	6 BD103687	BD103687 Myocardia

ALIGNMENTS

RESULT 1
AR380928
LOCUS AR380928 6263 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1473 from patent US 6607879.
ACCESSION AR380928
VERSION AR380928.1 GI:40088562
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 6263)
AUTHORS Cocks,B.G., Stuart,S.G. and Seilhamer,J.J.
TITLE Compositions for the detection of blood cell and immunological response gene expression
JOURNAL Patent: US 6607879-A 1473 19-AUG-2003;
FEATURES Location/Qualifiers
source 1..6263
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN	Query Match	100.0%	Score 6263	DB 6	Length 6263
	Best Local Similarity	100.0%	Pred. No. 0		
	Matches 6263	Conservative	0	Mismatches	0
			0	Indels	0
			0	Gaps	0
Qy	1	GCTAACTCAAGGAGACGCTCTGGTGAACACCCGCTGGGATCTAAAGAACCAAGCTCTGAAAG	60		
Db	1	GCTAACTCAAGGAGACGCTCTGGTGAACACCCGCTGGGATCTAAAGAACCAAGCTCTGAAAG	60		
Qy	61	TGTTCCAGCTGAATTTTCAGATCGGACAGACTCGCTCGGCTCCGGAGGAGTGAATCCA	120		
Db	61	TGTTCCAGCTGAATTTTCAGATCGGACAGACTCGCTCGGCTCCGGAGGAGTGAATCCA	120		
Qy	121	AGCTGCTCGCGACGCTGTCGCAAGTCGAGGATGTCGACCTAGCAGGCTGCTGCTG	180		
Db	121	AGCTGCTCGCGACGCTGTCGCAAGTCGAGGATGTCGACCTAGCAGGCTGCTGCTG	180		
Qy	181	CTGCTCTCTCACTTTTCTTCTCCCGACGAGTGTGAGACACTTCAAGGTTTACACGAACA	240		
Db	181	CTGCTCTCTCACTTTTCTTCTCCCGACGAGTGTGAGACACTTCAAGGTTTACACGAACA	240		
Qy	241	CCCGTTGATCAGACAGGGGTCTCTGGCGGAGTTCCTTTTCATCTGCCAAGCTACGGGA	300		

|||||
Db 241 CCCGTTGATCAGACAGGGGCTCTGGGGAGTTGCCCTCTTTCATCTGCCAGCTACGGGA 300
Qy 301 GACCAAGACCTAAATTTGTCGAAACAAAAGGAAAGAAAGTTCAGCAATTCAGATTT 360
Db 301 GACCAAGACCTAAATTTGTCGAAACAAAAGGAAAGAAAGTTCAGCAATTCAGATTT 360
Qy 361 GAGGTAATAGATTGAGCATGGGTCTGGATCGATCTCAGCAATTCAGCAATTCAGGACT 420
Db 361 GAGGTAATAGATTGAGCATGGGTCTGGATCGATCTCAGCAATTCAGCAATTCAGGACT 420
Qy 421 CCAGGAGATGAGGCATTTTAAATGTTGGCTTCAAAATGTTGGGAGAAATTAAGTGA 480
Db 421 CCAGGAGATGAGGCATTTTAAATGTTGGCTTCAAAATGTTGGGAGAAATTAAGTGA 480
Qy 481 TCCACCAAGACTCACAGTTTTCGCGGAAGATCAAAATCCAGGGGCTTCCCTACCAATTGAC 540
Db 481 TCCACCAAGACTCACAGTTTTCGCGGAAGATCAAAATCCAGGGGCTTCCCTACCAATTGAC 540
Qy 541 ATGGGCCCCACAGTTGAAGTGGTTGAGGCTACTCGCAGGGCCACCATGCTTTGTGACGC 600
Db 541 ATGGGCCCCACAGTTGAAGTGGTTGAGGCTACTCGCAGGGCCACCATGCTTTGTGACGC 600
Qy 601 AGTGTAATCCGGATCCAGAAATCACTTGGTTTAAAGATTTCTTACCTGTGGACACAAGC 660
Db 601 AGTGTAATCCGGATCCAGAAATCACTTGGTTTAAAGATTTCTTACCTGTGGACACAAGC 660
Qy 661 AACAAATGTCGTAATTAAGCAGTTACGATCAGAAATCTATTTGGTGGTACACCAATAAGA 720
Db 661 AACAAATGTCGTAATTAAGCAGTTACGATCAGAAATCTATTTGGTGGTACACCAATAAGA 720
Qy 721 GGAGCCCTTCAGATTGACAGTGAAGTCTGACCAAGGAAATATGAGTGTGTGCC 780
Db 721 GGAGCCCTTCAGATTGACAGTGAAGTCTGACCAAGGAAATATGAGTGTGTGCC 780
Qy 781 ACCAAGCGGGGCACTCGCTATTCGGCTTCTGCCAATTTATATGTACAGAGCTCGGA 840
Db 781 ACCAAGCGGGGCACTCGCTATTCGGCTTCTGCCAATTTATATGTACAGAGCTCGGA 840
Qy 841 GAAGTTCGCGTGTCCCAACCAAGATTCCTATCCCAACCACTAATCATGAATCATGCCA 900
Db 841 GAAGTTCGCGTGTCCCAACCAAGATTCCTATCCCAACCACTAATCATGAATCATGCCA 900
Qy 901 GGCGGAAGCGTTAATATCACTGTGTGGCGTGGGTCACCAATGCCCTTATGTAAGTGG 960
Db 901 GGCGGAAGCGTTAATATCACTGTGTGGCGTGGGTCACCAATGCCCTTATGTAAGTGG 960
Qy 961 ATGTTGGGGGCAAGATCTGACACCTGAAGATGATGCCAATAGGAAGAAATGTGCTA 1020
Db 961 ATGTTGGGGGCAAGATCTGACACCTGAAGATGATGCCAATAGGAAGAAATGTGCTA 1020
Qy 1021 GAACTGAATGATGAAGACAGTACGAAATTCACCTGTGTGTGTGTGTGTGTGTGTGTGT 1080
Db 1021 GAACTGAATGATGAAGACAGTACGAAATTCACCTGTGTGTGTGTGTGTGTGTGTGTGT 1080
Qy 1081 GTCAATGAGCAATAGCACAGATCACTGTCAAGCCTTACCAACCTCCAGGAATCCT 1140
Db 1081 GTCAATGAGCAATAGCACAGATCACTGTCAAGCCTTACCAACCTCCAGGAATCCT 1140
Qy 1141 GTAGTGACCGAGAGCAGACGATACAGCATCACCTGAGCTGGGACTCTGGGAACCTCGAG 1200
Db 1141 GTAGTGACCGAGAGCAGACGATACAGCATCACCTGAGCTGGGACTCTGGGAACCTCGAG 1200
Qy 1201 CCTGTTCTTATTAATTAATTCAGATTAACCTTAAATCTCTGAGGAATCTTCAAGAA 1260
Db 1201 CCTGTTCTTATTAATTAATTCAGATTAACCTTAAATCTCTGAGGAATCTTCAAGAA 1260
Qy 1261 ATTGATGGGGTGGCGACACAGCTACAGTGTGCTGAGTAAAGTCCCTACTCGGATTA 1320
Db 1261 ATTGATGGGGTGGCGACACAGCTACAGTGTGCTGAGTAAAGTCCCTACTCGGATTA 1320
Qy 1321 GAAATTCAGGGTGTGTGCTCAATTAACATTTGGGGGGGCTTCCAGCGAAACCTGTGCTA 1380

Db 1321 GAAATTCAGGGTGTGTGCTCAATTAACATTTGGGGGGGCTTCCAGCGAAACCTGTGCTA 1380
Qy 1381 ACACAAACCTCAGAGCAGACCATTCAGTGTGCCCGAGGGATGTCCAGGACAGAAATGTTG 1440
Db 1381 ACACAAACCTCAGAGCAGACCATTCAGTGTGCCCGAGGGATGTCCAGGACAGAAATGTTG 1440
Qy 1441 AGTTTCAGACCACTTTTGTGTAAGTGAAGAACTTGAAGAGCCAAATGACACATCCAA 1500
Db 1441 AGTTTCAGACCACTTTTGTGTAAGTGAAGAACTTGAAGAGCCAAATGACACATCCAA 1500
Qy 1501 GGATATAGAGTTTATTAATCAATGGATCCCACTCAACATGTCAACATGGATGAACAC 1560
Db 1501 GGATATAGAGTTTATTAATCAATGGATCCCACTCAACATGTCAACATGGATGAACAC 1560
Qy 1561 AATGTAGCTGACAGCCAAATCACTATTTGGCAACTAGTGTGCCCGAGAAACATATTTCT 1620
Db 1561 AATGTAGCTGACAGCCAAATCACTATTTGGCAACTAGTGTGCCCGAGAAACATATTTCT 1620
Qy 1621 GTCAAAAGTCTGTGCTTACCTCAATTTGGAGATGGTCCCTTTCAAGTGACATACAAAGTC 1680
Db 1621 GTCAAAAGTCTGTGCTTACCTCAATTTGGAGATGGTCCCTTTCAAGTGACATACAAAGTC 1680
Qy 1681 ATCACTCAGACAGAGTACCAAGGCGAGCCACTAAACTTCAAAAGCAGAACTGTGAA 1740
Db 1681 ATCACTCAGACAGAGTACCAAGGCGAGCCACTAAACTTCAAAAGCAGAACTGTGAA 1740
Qy 1741 ACAAGTATTTTGTCTCTTGGACACCTCCAGCTTCAGATACCATTTGCCAACTATGAAGTC 1800
Db 1741 ACAAGTATTTTGTCTCTTGGACACCTCCAGCTTCAGATACCATTTGCCAACTATGAAGTC 1800
Qy 1801 GTCTCAAAAGATGGGAGCATGGAGAGCAACGAATTAACCATTTGAGCCAGGACATCA 1860
Db 1801 GTCTCAAAAGATGGGAGCATGGAGAGCAACGAATTAACCATTTGAGCCAGGACATCA 1860
Qy 1861 TATAGGCTGCAAGACTGAAACCAAAACAGCTTATATTTTCGGTCTGGCTGCAAGCTCC 1920
Db 1861 TATAGGCTGCAAGACTGAAACCAAAACAGCTTATATTTTCGGTCTGGCTGCAAGCTCC 1920
Qy 1921 CCTCAAGGCTGGGTCTTCTACTGCAAGAAATACAGTACCAAGTTCAGTGTGCAAGGCG 1980
Db 1921 CCTCAAGGCTGGGTCTTCTACTGCAAGAAATACAGTACCAAGTTCAGTGTGCAAGGCG 1980
Qy 1981 TCAGCTCTCTCTCAAGACATTTAGTTCACAGCCAGCCAAAGTTCACATAGTATTTTGGTAA 2040
Db 1981 TCAGCTCTCTCTCAAGACATTTAGTTCACAGCCAGCCAAAGTTCACATAGTATTTTGGTAA 2040
Qy 2041 TGGCAACCTCCAAGTGGGAAACAGAAATGGAATGGAATTCATGAAATTCATCAAGTAC 2100
Db 2041 TGGCAACCTCCAAGTGGGAAACAGAAATGGAATGGAATTCATGAAATTCATCAAGTAC 2100
Qy 2101 ACTGCAAGTGGGAGGAGTACAGCCCTCAGAGATTTTGGGAATTTCTTCCGACACT 2160
Db 2101 ACTGCAAGTGGGAGGAGTACAGCCCTCAGAGATTTTGGGAATTTCTTCCGACACT 2160
Qy 2161 ACCAAATACCTTTTGGAAACAGCTGGGAAATGGAATGGAATTCAGTGTGCAAGCC 2220
Db 2161 ACCAAATACCTTTTGGAAACAGCTGGGAAATGGAATGGAATTCAGTGTGCAAGCC 2220
Qy 2221 CATACAGATGTGCGCCCTGGCCCTGAGAGCTGTGCTGTGTGATTCGAAACCAATGAAGAT 2280
Db 2221 CATACAGATGTGCGCCCTGGCCCTGAGAGCTGTGCTGTGTGATTCGAAACCAATGAAGAT 2280
Qy 2281 GTTCTTAGTGTGCTCTCTCCCAAGTGCAGTGTGAGGCTGTCAACTCAACATCTGTTAAA 2340
Db 2281 GTTCTTAGTGTGCTCTCTCCCAAGTGCAGTGTGAGGCTGTCAACTCAACATCTGTTAAA 2340
Qy 2341 GTCTCATGGGCTCACCCGTCGCAATTAACAGCATGCGCCAGATAAGAGGATATCAGGTG 2400
Db 2341 GTCTCATGGGCTCACCCGTCGCAATTAACAGCATGCGCCAGATAAGAGGATATCAGGTG 2400
Qy 2401 CATTTATGTAGGATGGAATAATGTGTAGCCCAAGGGCCAGCCCATGTCTGAAAGATGTCA 2460
Db 2401 CATTTATGTAGGATGGAATAATGTGTAGCCCAAGGGCCAGCCCATGTCTGAAAGATGTCA 2460

QY 2461 CTGGCTGATGCACAGTGGGAATTTGATGATCTACTGAAACATGACATGATCATTTCTGGG 2520
DB 2461 CTGGCTGATGCACAGTGGGAATTTGATGATCTACTGAAACATGACATGATCATTTCTGGG 2520
QY 2521 CTCACGCTGAAACTTCTACTCCCTCACCGTCAACAGCTTACAGGAGGAGATGCT 2580
DB 2521 CTCACGCTGAAACTTCTACTCCCTCACCGTCAACAGCTTACAGGAGGAGATGCT 2580
QY 2581 GCTCGCAGCAAGCCCAAACTGGTGTCCCACTGGGGGAGTTCCAGGGAAACCTCGGCTT 2640
DB 2581 GCTCGCAGCAAGCCCAAACTGGTGTCCCACTGGGGGAGTTCCAGGGAAACCTCGGCTT 2640
QY 2641 GTGATTAACCACTCAGATGATGATCTGCTTATTCAGTGGCACCCTCCGGTGCACACA 2700
DB 2641 GTGATTAACCACTCAGATGATGATCTGCTTATTCAGTGGCACCCTCCGGTGCACACA 2700
QY 2701 TTTGGACCTTTCAGGCTACCGTCTAAAAATTTGGCCGCAAGGATATGGAGCCACTTACT 2760
DB 2701 TTTGGACCTTTCAGGCTACCGTCTAAAAATTTGGCCGCAAGGATATGGAGCCACTTACT 2760
QY 2761 ACTCTGAGTCTTCTGAAAAAGAGATCACTTTACAGCTACAGACATCCACAGGGAGCA 2820
DB 2761 ACTCTGAGTCTTCTGAAAAAGAGATCACTTTACAGCTACAGACATCCACAGGGAGCA 2820
QY 2821 TCATACGCTTTCAGGCTCTCAGCCAGAAAAAAGTGGGCTTTGGGGAGGAGATGGTGAAG 2880
DB 2821 TCATACGCTTTCAGGCTCTCAGCCAGAAAAAAGTGGGCTTTGGGGAGGAGATGGTGAAG 2880
QY 2881 GAGATTTCCATTCAGAGAGAGTACCACTGATTCCTCAAAACCTTCACTCAGAGGCG 2940
DB 2881 GAGATTTCCATTCAGAGAGAGTACCACTGATTCCTCAAAACCTTCACTCAGAGGCG 2940
QY 2941 ACCACTTCAAACCTCGTCCAGTTATCTTGGCAACCACTGCTCTGGCAGAGAGAAATGGC 3000
DB 2941 ACCACTTCAAACCTCGTCCAGTTATCTTGGCAACCACTGCTCTGGCAGAGAGAAATGGC 3000
QY 3001 ATTATCACCAGTATACCTCTTTTATAGGATATCAACATCCCTCTCCCGATGGAG 3060
DB 3001 ATTATCACCAGTATACCTCTTTTATAGGATATCAACATCCCTCTCCCGATGGAG 3060
QY 3061 CAGCTTATTTGTCAGCTGACACCACTATGACACTGCTTAAACACCAAGTACCAACA 3120
DB 3061 CAGCTTATTTGTCAGCTGACACCACTATGACACTGCTTAAACACCAAGTACCAACA 3120
QY 3121 TAGCATGTAAGTACGTCTCATAAGCAAGGGCCCGGCCATATAGTCCAGTGTC 3180
DB 3121 TAGCATGTAAGTACGTCTCATAAGCAAGGGCCCGGCCATATAGTCCAGTGTC 3180
QY 3181 CAGTTTCCAGCACTGCTGTGGATCAAGTGTTCGCAAAATTTTCATGTCAAAAGCAGTA 3240
DB 3181 CAGTTTCCAGCACTGCTGTGGATCAAGTGTTCGCAAAATTTTCATGTCAAAAGCAGTA 3240
QY 3241 ATGAAGACTTCCTGTTGCTGTTGGAGATTCAGAGAAATTAATCTCCGCAATGCT 3300
DB 3241 ATGAAGACTTCCTGTTGCTGTTGGAGATTCAGAGAAATTAATCTCCGCAATGCT 3300
QY 3301 TTCAAAATTTCTTATGATGATGGGAAATGTTAGAGAAATTTTCATGTCAAAAGCAGTA 3360
DB 3301 TTCAAAATTTCTTATGATGATGGGAAATGTTAGAGAAATTTTCATGTCAAAAGCAGTA 3360
QY 3361 AAGTTAATTTGCACTGAGCAAGCTCAGGCAAGGATTCATATTTGCTGACAAATCGTGA 3420
DB 3361 AAGTTAATTTGCACTGAGCAAGCTCAGGCAAGGATTCATATTTGCTGACAAATCGTGA 3420
QY 3421 AACAGTGTGTGGCTGCACACAGGCTCAGGCAAGGATTCAGGCAAGGATTCATATTTGCT 3480
DB 3421 AACAGTGTGTGGCTGCACACAGGCTCAGGCAAGGATTCAGGCAAGGATTCATATTTGCT 3480
QY 3481 ACCAAGCTGCTTTCATTTGGGAAGACCACTTGGATGGCATGATTTACTGTGCAACTGCT 3540
DB 3481 ACCAAGCTGCTTTCATTTGGGAAGACCACTTGGATGGCATGATTTACTGTGCAACTGCT 3540

QY 3541 GAAGTACTGCAATGAGAAATATAAAGTTTACTACATAAATTTGTSCCTTTGAAGAAA 3600
DB 3541 GAAGTACTGCAATGAGAAATATAAAGTTTACTACATAAATTTGTSCCTTTGAAGAAA 3600
QY 3601 TCTCGGGGAAATTTATCAAGCCATGGAGAGTCCAGATGAAATGGAATTTAGATGAGCTG 3660
DB 3601 TCTCGGGGAAATTTATCAAGCCATGGAGAGTCCAGATGAAATTTAGATGAGCTG 3660
QY 3661 CTTAAGGAGATATCTAGGAGCGCAGAGCATCGTTATGGGAGAGAAATTTGAAATTAAG 3720
DB 3661 CTTAAGGAGATATCTAGGAGCGCAGAGCATCGTTATGGGAGAGAAATTTGAAATTAAG 3720
QY 3721 CCATATATTCGCTCACTTTGATGCTCTCCCACTGATTTCACTCGGGGAGTACAAAG 3780
DB 3721 CCATATATTCGCTCACTTTGATGCTCTCCCACTGATTTCACTCGGGGAGTACAAAG 3780
QY 3781 CATTATGCTGATTTACAAACAGCAACTCCAAAGTGGTCAAGAAATATGCTTCTTTCTG 3840
DB 3781 CATTATGCTGATTTACAAACAGCAACTCCAAAGTGGTCAAGAAATATGCTTCTTTCTG 3840
QY 3841 TTACAGTAAATGGAACATGACAGTCTAAGATGATGCAACAGCCCTTACTCCGACCCC 3900
DB 3841 TTACAGTAAATGGAACATGACAGTCTAAGATGATGCAACAGCCCTTACTCCGACCCC 3900
QY 3901 GTGCTGCTCAATGCTGATTCGGATCCGACCAATCAGGATGAAAGAGGCTTGTCTGG 3960
DB 3901 GTGCTGCTCAATGCTGATTCGGATCCGACCAATCAGGATGAAAGAGGCTTGTCTGG 3960
QY 3961 GTTGTAGCTCTGCTGCTTGCAGTGGTCTTTATCATCTGCTGCTTCTTCTTCTT 4020
DB 3961 GTTGTAGCTCTGCTGCTTGCAGTGGTCTTTATCATCTGCTGCTTCTTCTTCTT 4020
QY 4021 TATAAAGGAGAGGGCAGAGTCCGACTCTAGAAAAAGCAGCATACCGAAACAATAAGAG 4080
DB 4021 TATAAAGGAGAGGGCAGAGTCCGACTCTAGAAAAAGCAGCATACCGAAACAATAAGAG 4080
QY 4081 ATCCCTTTCACACCAACCAAGAGCTGTAGAACTGAGGGCCCTTACTTTCAACACCG 4140
DB 4081 ATCCCTTTCACACCAACCAAGAGCTGTAGAACTGAGGGCCCTTACTTTCAACACCG 4140
QY 4141 GGTATGCTAGCCATCTCCAAATACCCATCTTGAACTTGCAGACCACTTGAAGATG 4200
DB 4141 GGTATGCTAGCCATCTCCAAATACCCATCTTGAACTTGCAGACCACTTGAAGATG 4200
QY 4201 AAAGCAATGACAACTTGAAGTTTCCAGGAAATATGAGTCAATTTGACCCCTGGCCAG 4260
DB 4201 AAAGCAATGACAACTTGAAGTTTCCAGGAAATATGAGTCAATTTGACCCCTGGCCAG 4260
QY 4261 TTCACTTTGGGAACTTCAAACTTGAAGTAAACAAACCAAGATAGATACCGCAATGTA 4320
DB 4261 TTCACTTTGGGAACTTCAAACTTGAAGTAAACAAACCAAGATAGATACCGCAATGTA 4320
QY 4321 ATGCTATGATCATTTCCCGGTTCTCTATCAGCTATAGAGGGATCCCAAGGAGTAC 4380
DB 4321 ATGCTATGATCATTTCCCGGTTCTCTATCAGCTATAGAGGGATCCCAAGGAGTAC 4380
QY 4381 TATGTGAATGCCAATCATAGATGGTATAGGAAGCAAAATGCTATATTTGCAACACAG 4440
DB 4381 TATGTGAATGCCAATCATAGATGGTATAGGAAGCAAAATGCTATATTTGCAACACAG 4440
QY 4441 GGATCTCTCCCGGAACTTTGGGACTTTTGGGAAATGATATGGGAAACAGGAGTGC 4500
DB 4441 GGATCTCTCCCGGAACTTTGGGACTTTTGGGAAATGATATGGGAAACAGGAGTGC 4500
QY 4501 ACAGTGTCTATGATGACAAACTAGAGAAAGATCAAGGGTGAAGTGTGACCACTTGG 4560
DB 4501 ACAGTGTCTATGATGACAAACTAGAGAAAGATCAAGGGTGAAGTGTGACCACTTGG 4560
QY 4561 CCTAGCAGAGGCAAGAAACCCAGGACTCGTTCAAGTAAACGCTGCTTGTGAGAG 4620
DB 4561 CCTAGCAGAGGCAAGAAACCCAGGACTCGTTCAAGTAAACGCTGCTTGTGAGAG 4620
QY 4621 CTGGCCACATATTTGCTTGCACATTTGCAAGTGGTTCAGAGATGGTTCAGAGAGAG 4680

Db 4621 CTGCCACATATTGTGTTGCGAACAATTTGCACTTTTACAGAATGGTTCAAGTGAGAGAGA 4680
Qy 4681 GAAGTGAGACAAATCCAGTTTCAACCGCTCGGCTGATCATGGTGTTCAGAAACCCCTACA 4740
Db 4681 GAAGTGAGACAAATCCAGTTTCAACCGCTCGGCTGATCATGGTGTTCAGAAACCCCTACA 4740
Qy 4741 CCTTTCTAGCTTTCTTACGTAGAGTCAAAACCTGTAAACCTCCCGATGCTGGTCCGATG 4800
Db 4741 CCTTTCTAGCTTTCTTACGTAGAGTCAAAACCTGTAAACCTCCCGATGCTGGTCCGATG 4800
Qy 4801 GTTGTGCACTGCAGTGGGAGTTGGCGGACTGGTTGCTTCATCGTCATAGATGCCATG 4860
Db 4801 GTTGTGCACTGCAGTGGGAGTTGGCGGACTGGTTGCTTCATCGTCATAGATGCCATG 4860
Qy 4861 TTAGAAAGAAATAAGCATGAAAAACTGTAGATTAATTTATGGCCATGTAACTTTAATGAGA 4920
Db 4861 TTAGAAAGAAATAAGCATGAAAAACTGTAGATTAATTTATGGCCATGTAACTTTAATGAGA 4920
Qy 4921 GCCCAGAGGAACATATATGTTTCAACACAGAGACCAATACATCTTTATCCATGATGCACTG 4980
Db 4921 GCCCAGAGGAACATATATGTTTCAACACAGAGACCAATACATCTTTATCCATGATGCACTG 4980
Qy 4981 TTAGAACGAGTGAATTTGTGAAATATACCAGAGTGGCAGCTAGAAAATTGTATGCCCTACATT 5040
Db 4981 TTAGAACGAGTGAATTTGTGAAATATACCAGAGTGGCAGCTAGAAAATTGTATGCCCTACATT 5040
Qy 5041 CAGAGCTGACACAAATAGAAAACGGGAGAGAAATGTCAACAGAAATGGAGCTCGAATTTAAG 5100
Db 5041 CAGAGCTGACACAAATAGAAAACGGGAGAGAAATGTCAACAGAAATGGAGCTCGAATTTAAG 5100
Qy 5101 CGTCTAGCAGCTCAAAAGCTCACACCTCAAGGTTTATCAGTGCATCTTCCATGTAAT 5160
Db 5101 CGTCTAGCAGCTCAAAAGCTCACACCTCAAGGTTTATCAGTGCATCTTCCATGTAAT 5160
Qy 5161 AAATTTCAAAATCGCTTGTATTAATATATGCAATATGAATCCACAGGGTATGCTGCAG 5220
Db 5161 AAATTTCAAAATCGCTTGTATTAATATATGCAATATGAATCCACAGGGTATGCTGCAG 5220
Qy 5221 CCTATCGTGAGTAGAAGATCTGATTAATCAATGCCAGTTTATTTATGATGATACAGA 5280
Db 5221 CCTATCGTGAGTAGAAGATCTGATTAATCAATGCCAGTTTATTTATGATGATACAGA 5280
Qy 5281 CAACAGAAAGCTACATCGCTTACCAGGGCCCTTGGCAGAGACCTGGAAGCTCTGCG 5340
Db 5281 CAACAGAAAGCTACATCGCTTACCAGGGCCCTTGGCAGAGACCTGGAAGCTCTGCG 5340
Qy 5341 CGGATGCTCTGGGAACAAATTTCCACCATAGTTGTGATGCTCACCAAGCTCGGTGAAATG 5400
Db 5341 CGGATGCTCTGGGAACAAATTTCCACCATAGTTGTGATGCTCACCAAGCTCGGTGAAATG 5400
Qy 5401 GGCAGAGAAATGTCAACCAATATGCGCAGCAGAAAGCTGTGCAAGATACCACTACTTT 5460
Db 5401 GGCAGAGAAATGTCAACCAATATGCGCAGCAGAAAGCTGTGCAAGATACCACTACTTT 5460
Qy 5461 GTTGTAGATCCATGCTGAGTACAAATGCGCAGATATATCTTCAAGGAATTCAGAGTC 5520
Db 5461 GTTGTAGATCCATGCTGAGTACAAATGCGCAGATATATCTTCAAGGAATTCAGAGTC 5520
Qy 5521 ACAGATCCAGGAGCGGCCAGTCCCGAACAATAGGAGGATTCAGTTTCACTGACTGCCA 5580
Db 5521 ACAGATCCAGGAGCGGCCAGTCCCGAACAATAGGAGGATTCAGTTTCACTGACTGCCA 5580
Qy 5581 GAGCAAGAGTGCACAAAGTCCGGAGAGGATTTATTCATCTCATCGGCAAGTCCATAA 5640
Db 5581 GAGCAAGAGTGCACAAAGTCCGGAGAGGATTTATTCATCTCATCGGCAAGTCCATAA 5640
Qy 5641 ACAAAGAAAGTGTGGCAGAGTGGACCCATTTTCACTGCTCAGCGGGGGTGGGA 5700
Db 5641 ACAAAGAAAGTGTGGCAGAGTGGACCCATTTTCACTGCTCAGCGGGGGTGGGA 5700
Qy 5701 AGAATCGAGTCTTTCATAACCGCTAAGCATCTTTTGGAAAGAAATGAGATATGAAGAGTT 5760
Db 5701 AGAATCGAGTCTTTCATAACCGCTAAGCATCTTTTGGAAAGAAATGAGATATGAAGAGTT 5760

Db 5701 AGAATCGAGTCTTTCATAACCGCTAAGCATCTTTTGGAAAGAAATGAGATATGAAGAGTT 5760
Qy 5761 GTAGATATCTTCCAGACTGTCAAAATGTTAAGAACACAAACGACAGCTATGGTACAGACA 5820
Db 5761 GTAGATATCTTCCAGACTGTCAAAATGTTAAGAACACAAACGACAGCTATGGTACAGACA 5820
Qy 5821 GAGGATCAATATCAGTCTTCTATCGTCCGCACTAGAGTACCTGGCAGCTTTTGACCAC 5880
Db 5821 GAGGATCAATATCAGTCTTCTATCGTCCGCACTAGAGTACCTGGCAGCTTTTGACCAC 5880
Qy 5881 TATGCAACGTAGAAAACCCCTGACCCATCTTGGATTTTACTACAGGCCCTTCAATATCCA 5940
Db 5881 TATGCAACGTAGAAAACCCCTGACCCATCTTGGATTTTACTACAGGCCCTTCAATATCCA 5940
Qy 5941 TGGAGTCTCTTCTGAGCCATACAGGCGACTTGAAGAAGTCTTCTTAACCTCTAGCTAACA 6000
Db 5941 TGGAGTCTCTTCTGAGCCATACAGGCGACTTGAAGAAGTCTTCTTAACCTCTAGCTAACA 6000
Qy 6001 ACTACTTAGTGGGACTATTACACACAAAACAAATTAATAAACAATATTATCCAGGTGACC 6060
Db 6001 ACTACTTAGTGGGACTATTACACACAAAACAAATTAATAAACAATATTATCCAGGTGACC 6060
Qy 6061 AGAATCTTTTGACATCGCCCTTCCACCATCTGCTCTAATAAACAATTTTAGGGGCCA 6120
Db 6061 AGAATCTTTTGACATCGCCCTTCCACCATCTGCTCTAATAAACAATTTTAGGGGCCA 6120
Qy 6121 AGGGAGGGAATGTTTAAAGAAAGTCTTGAATTTAGTATTTAGTATTTAGTATTTAGTAT 6180
Db 6121 AGGGAGGGAATGTTTAAAGAAAGTCTTGAATTTAGTATTTAGTATTTAGTATTTAGTAT 6180
Qy 6181 TGCTGACCTGCTGCTTCAATTTCTAACTGTGTAAACTTTTTTAAACAAAATGTATCATTCG 6240
Db 6181 TGCTGACCTGCTGCTTCAATTTCTAACTGTGTAAACTTTTTTAAACAAAATGTATCATTCG 6240
Qy 6241 ATAAAGTGAATTTTAAAGAGTT 6263
Db 6241 ATAAAGTGAATTTTAAAGAGTT 6263

RESULT 2
HUMPTPD
LOCUS
DEFINITION Homo sapiens protein tyrosine phosphatase delta mRNA, complete cds.
ACCESSION L38929
VERSION L38929.1 GI:755652
KEYWORDS protein tyrosine phosphatase delta; transmembrane protein.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 6263)
AUTHORS Pulido,R., Krueger,N.X., Serra-Pages,C., Saito,H. and Streuli,M.
TITLE Molecular characterization of the human transmembrane protein-tyrosine phosphatase delta. Evidence for tissue-specific expression of alternative human transmembrane protein-tyrosine phosphatase delta isoforms
J. Biol. Chem. 270 (12), 6722-6728 (1995)

JOURNAL MEDLINE 95204468
FUBMED 7896816
COMMENT Original
FEATURES
source
Location/Qualifiers
1. .6263
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1. .6263
/codon_start=1
/product="protein tyrosine phosphatase delta"
/protein_id="AAC41749.1"
/db_xref="GI:755653"
translation="MVHVKALLLLLLTFFLRDTAETPPRFRTRTPVDOTGSGGVASF1
COATGPRPKIVNKKGVNSQRFVIEFDGSGSVLRIOPLRTPRDEAIYECASN

NVGEISVSTRLTVLREDQIPRGFPPTIDMGPOLKVVVERTRTATMLCAASGNPDPEITWF
 KDFLPVDSNNNGRIKQURSESIGTPIRGALQIEQSSESQKTECVATNSAGTRY5
 APANLYRELREVRVPRFSPPIPTNHEIMPGSVNITCVAGSPMPYVXWMLGABDL
 TPEDDMPTGRNLENDVYRQANVYTCVAMSTLGVIEALQATVTRALPPPGTPVVTES
 TATSITLTWDSNPPSPVYIIIOHKPKNSELYKEIDGVATTVTRSVAGLSYSPSYDEFR
 VVANNIGRGPSPBVLQTSOAPSSAPRDVQARMLSTTILVQWKEPBPNGOIQ
 YRVVYTMPTOHOVNNMKNHVDASQITIGNLPQKTYSVKVLAFSTIGDPSLDIQ
 VITQGVGPQPLNFKABESETSIILSWTPRSDTIANYELYKDGEGESQRIIEP
 SILWSQGLKNSLYIFRLAASPOGLGASTAEISARTWQSPSPAPPODISCTSPST
 STLSWQPPPEKNGIITYESIKYTAVDGEDDKPHEILGIPSDTTKLLQLEKWE
 YRIITATHTDGLGPPESLVIRITNEDVPSGPRKVEAVNSTMVSKVSPSPVKNQ
 HGOIRGVHYVYRMENGSPKQMLDVLADAOWEFDOTTEHDMIISGLQPSYSYL
 TVTA VTKDGARSPLKSVITGAVGKPRLVINHOMNTALIOHPPEVMDTPGLOGY
 BLKFRKDMPLTLEPSEDEHFTATDHIKGSVYFRLSARNKVGFGEEMKESIP
 BEVPTQFQNLHSEGTTSVOLSQWPPTDPAERNGIITKTYLLYRDINILPLMEQLI
 VPADTMTLTGKDPDTYDVYVRAHSTSGPGYSPSQVQFTLPVDVFPVFKVAM
 KTSVLLSWEIIPENTNSAMPFKILYDDGKQVEVDGRATQKLIIVLKPESKYSFVLNVR
 GNSAGKGPITKWPESDEMEDELLEKEISRKERSIRYGREVELKPYIAHPDVLTEFT
 LKKSRGKFIKWPESDEMEDELLEKEISRKERSIRYGREVELKPYIAHPDVLTEFT
 LGDDKHGFGTNKQLQSQEYVFFVLAWEHASKMYATSPYSDPVMSMDLPQITD
 EEEGLIIVVGPVLA VVPIICIVIALILYKRAESDSRKSIPNNKEIPSHHPTDVE
 LRRLNFOTGPHHPPIPILEBLADHIERLKANDNLKFSQYESIDPGQFTWESNLE
 VNKPNRYIAVADHSVLLSAIEGIPGSDYVNAVYIDGYRKQNAVITATQGLSPETP
 GDFWRMIWEORSATVMMTKLEERSRVKCDYWPSSRGTEHGLVQVTLTDLTVELATVC
 VRTALYKNGSSEKEVRFQFTAPDQHGVEHPPTFLAFLRKYTCNPPDAGPMVH
 CSAGVGTGCTIVIDAMLERIKHEKTVDIVGHVTLMRQNRNMVQTEQYIFIDHALL
 EAVTCNGTEVPARMLIATYIKLTQIETGENTVGMELFPRKLASKAHSTRFISANLPC
 NFKNRLNIMPYESTVCLQPIRGVEGSDYINASFIDGYRQKAYATATQGLPAAETL
 DFWMLWHEHNSITIVMMLTKREMREKCHOYMPAERSARYQYFVVDPMARVMPQYIE
 REFVKTVARDQSKTRVQFTPDWPEQGVKSGRGFDIFIGOVHKTKEQFGDQGPISV
 HCSAGVGTGCTVITLSVLERMYVEGVVDIFQTVKMLRTQRPANVQTEQYQFSYRAA
 LEYLGSEFPHYAT"

6263

polyA_site
ORIGIN

Query Match	100.0%;	Score 6263;	DB 9;	Length 6263;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 6263;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	1	GCTAACTCAAGGAGAGCGTCTGGTGAACACCCGTTGGGATCTTAAAGAACAAAGCTCTGAAAG	60		
Db	1	GCTAACTCAAGGAGAGCGTCTGGTGAACACCCGTTGGGATCTTAAAGAACAAAGCTCTGAAAG	60		
Qy	61	TGTTCCAGCTCAAAATTTTCAGATCGGACAGACTCGTGGCGGCTCGGAGGCGAGTATCCA	120		
Db	61	TGTTCCAGCTCAAAATTTTCAGATCGGACAGACTCGTGGCGGCTCGGAGGCGAGTATCCA	120		
Qy	121	AGCTGCTCGGCACGCTGCTCCAGCTGACAGGATGGTGACAGTACGAGGCTCGCTGCTG	180		
Db	121	AGCTGCTCGGCACGCTGCTCCAGCTGACAGGATGGTGACAGTACGAGGCTCGCTGCTG	180		
Qy	181	CTGCTCTCTACTTTTCTCCGACCGGATGCTGAGACACCTCCAAAGGTTTACACGAACA	240		
Db	181	CTGCTCTCTACTTTTCTCCGACCGGATGCTGAGACACCTCCAAAGGTTTACACGAACA	240		
Qy	241	CCCGTTGATCAGACAGGGGCTCTGGCGGAGTTCGCTCTTTCATCTGCGAAGCTACGGGA	300		
Db	241	CCCGTTGATCAGACAGGGGCTCTGGCGGAGTTCGCTCTTTCATCTGCGAAGCTACGGGA	300		
Qy	301	GACCAAGAGCTCAAAATTTGCTCGGAAACAAAAGGAAAGAAAGTTCAGCAATCAGAGATTT	360		
Db	301	GACCAAGAGCTCAAAATTTGCTCGGAAACAAAAGGAAAGAAAGTTCAGCAATCAGAGATTT	360		
Qy	361	GAGTAAATAGAGTTTGAAGTGGGTCTGGATCAGTCTTCAGAAATACAAACCTTACGGACT	420		
Db	361	GAGTAAATAGAGTTTGAAGTGGGTCTGGATCAGTCTTCAGAAATACAAACCTTACGGACT	420		
Qy	421	CCGAGGATGAGGCCATTTATGAATGTGTGGCTCAAAATTAATGTGGGAGAAATTAAGTGA	480		
Db	421	CCGAGGATGAGGCCATTTATGAATGTGTGGCTCAAAATTAATGTGGGAGAAATTAAGTGA	480		
Qy	481	TCCACAGAGCTCAGATTTTGGGGGAGAGTCAAAATTCACAGGGGCTTCCCTACCATTTGAC	540		
Db	481	TCCACAGAGCTCAGATTTTGGGGGAGAGTCAAAATTCACAGGGGCTTCCCTACCATTTGAC	540		

Db	481	TCCACAGAGCTCAGAGTTTTCGGGGAAGATCAAAATTTCCAGGGGCTTCCCTACCATTTGAC	540		
Qy	541	ATGGGCCACAGTTGAAGGTGGTGGAGCGTACTCGCAGCGCCACCATGCTTTGTGCAGCC	600		
Db	541	ATGGGCCACAGTTGAAGGTGGTGGAGCGTACTCGCAGCGCCACCATGCTTTGTGCAGCC	600		
Qy	601	AGTGGTAATCCGGATCCAGAAATCAGTGGTTTAAAGATTTCTTACCTGTGGACACAAGC	660		
Db	601	AGTGGTAATCCGGATCCAGAAATCAGTGGTTTAAAGATTTCTTACCTGTGGACACAAGC	660		
Qy	661	AACAACATGTGCTGATTTAAGCAGTTAGATCAGAACTATTTGGTGGTACACCAATAAGA	720		
Db	661	AACAACATGTGCTGATTTAAGCAGTTAGATCAGAACTATTTGGTGGTACACCAATAAGA	720		
Qy	721	GGAGCCCTTCAGATTGAGCAGAGTGAAGAGTCTGACCAAGGAAATATGAGTGTGTGCC	780		
Db	721	GGAGCCCTTCAGATTGAGCAGAGTGAAGAGTCTGACCAAGGAAATATGAGTGTGTGCC	780		
Qy	781	ACCAAGACGGGGGCACTCGCTATTTCGCTCTCTGCCAATTTATATGTCAGAGAGCTGGGA	840		
Db	781	ACCAAGACGGGGGCACTCGCTATTTCGCTCTCTGCCAATTTATATGTCAGAGAGCTGGGA	840		
Qy	841	GAGTTCGCGGTGCCCAACCAAGATTCTCTATCCACCACCTAATCATGAATCATGCCA	900		
Db	841	GAGTTCGCGGTGCCCAACCAAGATTCTCTATCCACCACCTAATCATGAATCATGCCA	900		
Qy	901	GGCGAAGCGTTAATATCACTGTGTGGCGGTGGGGTCAACCAATGCTTATGTAAGTGG	960		
Db	901	GGCGAAGCGTTAATATCACTGTGTGGCGGTGGGGTCAACCAATGCTTATGTAAGTGG	960		
Qy	961	ATGTTGGGGGAGAGATCTGACACCTGAAGATGATATGCCAATAGGAAGAAATGTGCTA	1020		
Db	961	ATGTTGGGGGAGAGATCTGACACCTGAAGATGATATGCCAATAGGAAGAAATGTGCTA	1020		
Qy	1021	GAACTGAATGATGTAAGCAGTCAAGCAATTTACACCTGTGTGTGTATGTCAACACTGGT	1080		
Db	1021	GAACTGAATGATGTAAGCAGTCAAGCAATTTACACCTGTGTGTGTATGTCAACACTGGT	1080		
Qy	1081	GTCAATTGAAGCAATAGCAGATCAGTGTCAAAAGCCTTACCACCACTCCAGGAACTCCT	1140		
Db	1081	GTCAATTGAAGCAATAGCAGATCAGTGTCAAAAGCCTTACCACCACTCCAGGAACTCCT	1140		
Qy	1141	GTAGTACCCAGAGCAGCTACAGCATCACATGACGTGGGACTCTGGGAAACCTGTAG	1200		
Db	1141	GTAGTACCCAGAGCAGCTACAGCATCACATGACGTGGGACTCTGGGAAACCTGTAG	1200		
Qy	1201	CTGTTTCTTTTATACATAAATTCAGCATAAAACCTTAAAGCTCTGAGGAACTTTACAAAGAA	1260		
Db	1201	CTGTTTCTTTTATACATAAATTCAGCATAAAACCTTAAAGCTCTGAGGAACTTTACAAAGAA	1260		
Qy	1261	ATTGATGGGTGGGACCAACACGCTACAGTGTGCTGGACTTAAGTCCCTACTCGGATTTAT	1320		
Db	1261	ATTGATGGGTGGGACCAACACGCTACAGTGTGCTGGACTTAAGTCCCTACTCGGATTTAT	1320		
Qy	1321	GAATTCAGGGTGTGCTGTCAATAAATTCAGTGGGGGGGGCTCCACAGGAACTCTGTGCTA	1380		
Db	1321	GAATTCAGGGTGTGCTGTCAATAAATTCAGTGGGGGGGGCTCCACAGGAACTCTGTGCTA	1380		
Qy	1381	ACAAACCTTCAGAGCAAGCACCATCCAGTCCCGGAGGATGTCCAGGGCAGCAATGTTG	1440		
Db	1381	ACAAACCTTCAGAGCAAGCACCATCCAGTCCCGGAGGATGTCCAGGGCAGCAATGTTG	1440		
Qy	1441	AGTTCGACCAACCAATTTTGGTACAGTGGAGGAACTTGAAGAGCAATAGGACAGATCCAA	1500		
Db	1441	AGTTCGACCAACCAATTTTGGTACAGTGGAGGAACTTGAAGAGCAATAGGACAGATCCAA	1500		
Qy	1501	GGATATAGAGTTTATTTATACATGGATCCCACTCAACATGTCAACAACTGGATGAAACAC	1560		
Db	1501	GGATATAGAGTTTATTTATACATGGATCCCACTCAACATGTCAACAACTGGATGAAACAC	1560		
Qy	1561	AATGTAGCTGACAGCCAAATCACTACTATTGGGCAATTTAGTGTCCCGGAGAAACATTTCT	1620		
Db	1561	AATGTAGCTGACAGCCAAATCACTACTATTGGGCAATTTAGTGTCCCGGAGAAACATTTCT	1620		

QY	1621	GTCAAAGTCTCGCTTTTACCTCAAATTGGAGATGGTCCCTTTCAAGTGACATACAAGTC	1680
DB	1621	GTCAAAGTCTCGCTTTTACCTCAAATTGGAGATGGTCCCTTTCAAGTGACATACAAGTC	1680
QY	1681	ATCACTCAGACAGGAGTACAGGGCAGCCACTAAACTTTCAAAGCAGAACTGAGTCTGAA	1740
DB	1681	ATCACTCAGACAGGAGTACAGGGCAGCCACTAAACTTTCAAAGCAGAACTGAGTCTGAA	1740
QY	1741	ACAAGTATTTTGCTCTCTTGGACACCTCCACGTTTCAGATACCAATGCCCACTATGAAC	1800
DB	1741	ACAAGTATTTTGCTCTCTTGGACACCTCCACGTTTCAGATACCAATGCCCACTATGAAC	1800
QY	1801	GTCTACAAAGATGGGAGCATGGAGAGCAGCAAGAAATACCAATTGAGCCAGGACATCA	1860
DB	1801	GTCTACAAAGATGGGAGCATGGAGAGCAGCAAGAAATACCAATTGAGCCAGGACATCA	1860
QY	1861	TATAGGCTGCAAGGACTGAAACCAAACAGCTTTATATCTATTTCCGTCTGGCTGCAC	1920
DB	1861	TATAGGCTGCAAGGACTGAAACCAAACAGCTTTATATCTATTTCCGTCTGGCTGCAC	1920
QY	1921	CCTCAAGGCTGGGTGCTTCTACTGCAAGAAATATCAGCTAGAACCAATGCAGTCAAA	1980
DB	1921	CCTCAAGGCTGGGTGCTTCTACTGCAAGAAATATCAGCTAGAACCAATGCAGTCAAA	1980
QY	1981	TCAGCTCTCTCTCAAGACATATTAGTTGACACAGCCCAAGTTCACCTAGTATTTTGGT	2040
DB	1981	TCAGCTCTCTCTCAAGACATATTAGTTGACACAGCCCAAGTTCACCTAGTATTTTGGT	2040
QY	2041	TGGCAACCTCCACAGTGGAAAAACAGAAATGCAATATATCACTGAATCTCCATCAAGT	2100
DB	2041	TGGCAACCTCCACAGTGGAAAAACAGAAATGCAATATATCACTGAATCTCCATCAAGT	2100
QY	2101	ACTGCAGTGGATGGGGAAGATGACAGACCTCAAGAGATTTTGGGAATTCCTTCGGAC	2160
DB	2101	ACTGCAGTGGATGGGGAAGATGACAGACCTCAAGAGATTTTGGGAATTCCTTCGGAC	2160
QY	2161	ACCAAAATACCTTTTGGAAACAGCTGGAAAAATGCAATGCAATACCGGATCACTG	2220
DB	2161	ACCAAAATACCTTTTGGAAACAGCTGGAAAAATGCAATGCAATACCGGATCACTG	2220
QY	2221	CATACAGATGTGGCCCTGGCCCTGAGAGCTGTGCTGGTTGATTCGAACCAATGAAG	2280
DB	2221	CATACAGATGTGGCCCTGGCCCTGAGAGCTGTGCTGGTTGATTCGAACCAATGAAG	2280
QY	2281	GTTCTAGTGTCTCTCTGCAAAAGTCGAGGTAGAGCTGTCAACTCAACATCTGTTAA	2340
DB	2281	GTTCTAGTGTCTCTCTGCAAAAGTCGAGGTAGAGCTGTCAACTCAACATCTGTTAA	2340
QY	2341	GTTCTATGCGGCTCACCGTGCCCAATAACAGCATGGCCAGATAGAGGATATCAGGT	2400
DB	2341	GTTCTATGCGGCTCACCGTGCCCAATAACAGCATGGCCAGATAGAGGATATCAGGT	2400
QY	2401	CAATTATGTAGGATGGAAAAATGCTGAGCCCAAGGGCCAGCCCATGCTGAAAGATG	2460
DB	2401	CAATTATGTAGGATGGAAAAATGCTGAGCCCAAGGGCCAGCCCATGCTGAAAGATG	2460
QY	2461	CTGGCTGATGACAGTGGGAATTTGATGATATCTGAACATGACATGATTCATTTCTG	2520
DB	2461	CTGGCTGATGACAGTGGGAATTTGATGATATCTGAACATGACATGATTCATTTCTG	2520
QY	2521	CTCCAGCTGAACTTCTCTACTCCCTCACCGTCAACAGCTTACACCAAGGAGATGGT	2580
DB	2521	CTCCAGCTGAACTTCTCTACTCCCTCACCGTCAACAGCTTACACCAAGGAGATGGT	2580
QY	2581	GCTCGCAGAACCCAACTGGTGTCAACCACTGGGCGAGTTCCAGGGAAACCTCGGCTT	2640
DB	2581	GCTCGCAGAACCCAACTGGTGTCCACCACTGGGCGAGTTCCAGGGAAACCTCGGCTT	2640
QY	2641	GTGATTAAACCACTCAGATGAATATCTGCTCTTATTCAGTGGCACCTCCGGTGGACA	2700
DB	2641	GTGATTAAACCACTCAGATGAATATCTGCTCTTATTCAGTGGCACCTCCGGTGGACA	2700

Qy	2701	TTTTGGACCTCTTCAGGGCTACCGTCTAAATAATTTGGCGCGCAAGATATGGAGCCACTTACT	2760
Db	2701	TTTTGGACCTCTTCAGGGCTACCGTCTAAATAATTTGGCGCGCAAGATATGGAGCCACTTACT	2760
Qy	2761	ACTCTTGAGTTCTCTGAAAAGAGATCACTTTACAGCTACACATCCACAAGGAGCA	2820
Db	2761	ACTCTTGAGTTCTCTGAAAAGAGATCACTTTACAGCTACACATCCACAAGGAGCA	2820
Qy	2821	TCATAGCTCTTCAGGCTCTCAGCCAGAAAACAAAGTGGCTTTGGGAGGAGATGTTGAAG	2880
Db	2821	TCATAGCTCTTCAGGCTCTCAGCCAGAAAACAAAGTGGCTTTGGGAGGAGATGTTGAAG	2880
Qy	2881	GAGATTTCCATTCAGAGAAGTACCAACTGGATTCCTCAAACTTCTACTCAGAAGGC	2940
Db	2881	GAGATTTCCATTCAGAGAAGTACCAACTGGATTCCTCAAACTTCTACTCAGAAGGC	2940
Qy	2941	ACCACITCAACCTCCGTCAGTTATCTTGCAACCACTGTCTCTGCGACAGAGAAATGCC	3000
Db	2941	ACCACITCAACCTCCGTCAGTTATCTTGCAACCACTGTCTCTGCGACAGAGAAATGCC	3000
Qy	3001	ATTATCACCAAGTATACCCCTCTTTATAGGATATCAACATCCCCCTTCTCCCGATGGAG	3060
Db	3001	ATTATCACCAAGTATACCCCTCTTTATAGGATATCAACATCCCCCTTCTCCCGATGGAG	3060
Qy	3061	CAGCTTATGTTCCAGCTGACACCACTATGACACTCACTGGCTTTAAAAACAGATATCCACA	3120
Db	3061	CAGCTTATGTTCCAGCTGACACCACTATGACACTCACTGGCTTTAAAAACAGATATCCACA	3120
Qy	3121	TACAGTGTAAAAGTACGTGCTCATACGAGCAAAAGGCGCGGGCCATATAGTCCAGTGTC	3180
Db	3121	TACGATGTAAAAGTACGTGCTCATACGAGCAAAAGGCGCGGGCCATATAGTCCAGTGTC	3180
Qy	3181	CAGTTACAGGACACTGCCCTGTGGATCAAGTCTTTGCAAAAATTTTCATGTCAAAGCAGTA	3240
Db	3181	CAGTTACAGGACACTGCCCTGTGGATCAAGTCTTTGCAAAAATTTTCATGTCAAAGCAGTA	3240
Qy	3241	ATGAAGACTTCCGTTGTGCTCTTGGAGATTCAGAGAAATTAACCTCCGCCATGCGCT	3300
Db	3241	ATGAAGACTTCCGTTGTGCTCTTGGAGATTCAGAGAAATTAACCTCCGCCATGCGCT	3300
Qy	3301	TTCAAAATTTCTTATGATGATGGAAAAATGGTAGAAGAGTGTGATGGCGAGCCACACAG	3360
Db	3301	TTCAAAATTTCTTATGATGATGGAAAAATGGTAGAAGAGTGTGATGGCGAGCCACACAG	3360
Qy	3361	AAGTTAATGTCAACCTGAACCTGTGAATCATATTTCTGTCTGCAAAATTCGTGGA	3420
Db	3361	AAGTTAATGTCAACCTGAACCTGTGAATCATATTTCTGTCTGCAAAATTCGTGGA	3420
Qy	3421	AACAGTCTCTGTGGCTGACAGCAGGGTCAACGGCAAGACTGCACACAGATGTTAAGT	3480
Db	3421	AACAGTCTCTGTGGCTGACAGCAGGGTCAACGGCAAGACTGCACACAGATGTTAAGT	3480
Qy	3481	ACCAAGCTCTGCTTCATTTGGGAAGACAACCTTGGATGGCATGATTAATGTCGCAACTGCT	3540
Db	3481	ACCAAGCTCTGCTTCATTTGGGAAGACAACCTTGGATGGCATGATTAATGTCGCAACTGCT	3540
Qy	3541	GAAGTACTGCAATGAGATATAAAGGTTACTACATAATAATTTGTCCTTTGAAGAAA	3600
Db	3541	GAAGTACTGCAATGAGATATAAAGGTTACTACATAATAATTTGTCCTTTGAAGAAA	3600
Qy	3601	TCTCGCGGAAATTTATCAAGCCATGGGAGAGTCCAGATGAAATGGAATTAGATGAGCTG	3660
Db	3601	TCTCGCGGAAATTTATCAAGCCATGGGAGAGTCCAGATGAAATGGAATTAGATGAGCTG	3660
Qy	3661	CTTAAGGAGATATCTAGGAAGCGCAGAACATTCGTTATCGGAGAGAGTTGAATTAAG	3720
Db	3661	CTTAAGGAGATATCTAGGAAGCGCAGAACATTCGTTATCGGAGAGAGTTGAATTAAG	3720
Qy	3721	CCATATATTGCCGTCACTTTGATGTCCTTCCCACTGAGTTCACTCCCTGGGGGATGACAAG	3780
Db	3721	CCATATATTGCCGTCACTTTGATGTCCTTCCCACTGAGTTCACTCCCTGGGGGATGACAAG	3780
Qy	3781	CATTATGGTGGATTTACAAAACAAGCAACTCCAAAGTGGTCAAGAAATATGCTCTTTGTG	3840

[illegible]

Db	4861	TTAGAAAGAAATAAAGCATGAAAAAATCTGTAGATATTTATGGCCATGTAACTTTTAATGAGA	4920
Qy	4921	GCCACAGAGAACTATATGGTTCAAAACAGAAAGACCAATACATCTTTATCTCATGTGCACTG	4980
Db	4921	GCCACAGAGAACTATATGGTTCAAAACAGAAAGACCAATACATCTTTATCTCATGTGCACTG	4980
Qy	4981	TTAGAAAGCAGTGACTGTGTGGAAATAACGAAAGTCGCAGCTAGAACTTTGTATGCTTACATTT	5040
Db	4981	TTAGAAAGCAGTGACTGTGTGGAAATAACGAAAGTCGCAGCTAGAACTTTGTATGCTTACATTT	5040
Qy	5041	CAGAACTGCACACAAATAGAAAACGGGAGAGAAATCTCACAGGAATGGAGCTCGAAATTTAAG	5100
Db	5041	CAGAACTGCACACAAATAGAAAACGGGAGAGAAATCTCACAGGAATGGAGCTCGAAATTTAAG	5100
Qy	5101	CGTCTAGCCAGCTCAAAAGCTCAACACCTCAAGGTTTTATCAGTGCCTAATCTTCCATGTAAAT	5160
Db	5101	CGTCTAGCCAGCTCAAAAGCTCAACACCTCAAGGTTTTATCAGTGCCTAATCTTCCATGTAAAT	5160
Qy	5161	AAATTTCAAAATCGCCTTGTTAATATATATGCCATATGAATCTCAAGGGTATGCCTGTCAG	5220
Db	5161	AAATTTCAAAATCGCCTTGTTAATATATATGCCATATGAATCTCAAGGGTATGCCTGTCAG	5220
Qy	5221	CCTATCCGTGGAGTAGAAGGATCTGATTAATCAATCCAGCTCAAGTTTATTTAGTGATGATACAGA	5280
Db	5221	CCTATCCGTGGAGTAGAAGGATCTGATTAATCAATCCAGCTCAAGTTTATTTAGTGATGATACAGA	5280
Qy	5281	CAACAGAAAGCTACATCGCTACCCAGGGGCCCTTGGCAGAGACCACTGAAGACTTCTCTGG	5340
Db	5281	CAACAGAAAGCTACATCGCTACCCAGGGGCCCTTGGCAGAGACCACTGAAGACTTCTCTGG	5340
Qy	5341	CGGATGCTCTGGGAAACAATTTCCACCATATGTTGTGATGCTCTACCAAGCTGCGTGAATG	5400
Db	5341	CGGATGCTCTGGGAAACAATTTCCACCATATGTTGTGATGCTCTACCAAGCTGCGTGAATG	5400
Qy	5401	GGCAGAGAGAAATGTTCAACCAATACTGGCCAGCAGAAACGGTCTCCAGAGTACCAAGTACTTT	5460
Db	5401	GGCAGAGAGAAATGTTCAACCAATACTGGCCAGCAGAAACGGTCTCCAGAGTACCAAGTACTTT	5460
Qy	5461	GTTCTAGATCCATGGCTGAGTACAACTGCCACAGTATATCTTAAGGGAAATTCAAAGTTC	5520
Db	5461	GTTCTAGATCCATGGCTGAGTACAACTGCCACAGTATATCTTAAGGGAAATTCAAAGTTC	5520
Qy	5521	ACAGATGCCAGGACGGCCAGTCCGAAACAGTAAAGCAGTTCCAGTCTCACTGACTGGCCA	5580
Db	5521	ACAGATGCCAGGACGGCCAGTCCGAAACAGTAAAGCAGTTCCAGTCTCACTGACTGGCCA	5580
Qy	5581	GAGCAAGAGTGCCAAAGTCCGAGAAAGGATTTATTTGACTTCATCGGCCCAAGTCCATAA	5640
Db	5581	GAGCAAGAGTGCCAAAGTCCGAGAAAGGATTTATTTGACTTCATCGGCCCAAGTCCATAA	5640
Qy	5641	ACAAAAGAACAGTTTGGCCAAAGATGGAACCATTTCACTGATTCAGTCAGCGCGGCGTGG	5700
Db	5641	ACAAAAGAACAGTTTGGCCAAAGATGGAACCATTTCACTGATTCAGTCAGCGCGGCGTGG	5700
Qy	5701	AGAACTGGAGTCTTCATAACGCTAAGCATTTGTTGGAAAGAAATGAGATATGAAGGAGTT	5760
Db	5701	AGAACTGGAGTCTTCATAACGCTAAGCATTTGTTGGAAAGAAATGAGATATGAAGGAGTT	5760
Qy	5761	GTAGATATCTTCAGACTGTCAAAATTTGAAGAACACACGACAGCTATGGTACAGACA	5820
Db	5761	GTAGATATCTTCAGACTGTCAAAATTTGAAGAACACACGACAGCTATGGTACAGACA	5820
Qy	5821	GAGGATCAATATCAGTTTTCTTATCGTGGCAGCTAGAGTACTCTGGGAGCTTTGACACC	5880
Db	5821	GAGGATCAATATCAGTTTTCTTATCGTGGCAGCTAGAGTACTCTGGGAGCTTTGACACC	5880
Qy	5881	TATGCAACGTAGAAAACCCCTGACCCCATTTCTGGATTTTTTACTACAGGCCCTTCAATATCA	5940
Db	5881	TATGCAACGTAGAAAACCCCTGACCCCATTTCTGGATTTTTTACTACAGGCCCTTCAATATCA	5940
Qy	5941	TGGAGTCTCTTCGAGCCATACAGGGCACTTGAGAAAGTCTTCTTTAACTTCTAGCTTAA	6000
Db	5941	TGGAGTCTCTTCGAGCCATACAGGGCACTTGAGAAAGTCTTCTTTAACTTCTAGCTTAA	6000

```
Qy 6001 ACTACTTAGTGGGACTATTACACACAAACAAATTAATAAACAATTAATTCAGGTGACC 6060
Db 6001 ACTACTTAGTGGGACTATTACACACAAACAAATTAATAAACAATTAATTCAGGTGACC 6060
Qy 6061 AAGAAATCTTTGACATCGCCCTTCCACCATCTGCTCATATAAACAATTTAGGGGCCA 6120
Db 6061 AAGAAATCTTTGACATCGCCCTTCCACCATCTGCTCATATAAACAATTTAGGGGCCA 6120
Qy 6121 AGGGAGGGAATGTTTAAAGAAAGTCTTGAATTTAGTTTATTTAGTTTATTTAGTAAATAC 6180
Db 6121 AGGGAGGGAATGTTTAAAGAAAGTCTTGAATTTAGTTTATTTAGTTTATTTAGTAAATAC 6180
Qy 6181 TGCTGACCTGTCTTCAATCTTAATCTGTAATCTTTTAAACAATTTATCATTCG 6240
Db 6181 TGCTGACCTGTCTTCAATCTTAATCTGTAATCTTTTAAACAATTTATCATTCG 6240
Qy 6241 ATAAAGTGAATTTTAAAGAAAGTT 6263
Db 6241 ATAAAGTGAATTTTAAAGAAAGTT 6263

RESULT 3
LOCUS CQ722829 5754 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 8763 from Patent WO02068579.
ACCESSION CQ722829
VERSION CQ722829.1 GI:42283686
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kites, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 8763 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source
1..5754
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="caxon:9606"

ORIGIN
Query Match 90.3%; Score 5657.6; DB 6; Length 5754;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 5731; Conservative 0; Mismatches 14; Indels 15; Gaps 5;

Qy 505 GAAGATCAAAATTCAGGGGCTTCCCTACCATTTGACATGGGCCACACATTTGAAGTGT 564
Db 9 GAAGATCAAAATTCAGGGGCTTCCCTACCATTTGACATGGGCCACACATTTGAAGTGT 68
Qy 565 GAGCTACTCCGACGGCCACCATGTTTTCGACGCTGTTTCCGATCCGATCCGATCCGATCC 624
Db 69 GAGCTACTCCGACGGCCACCATGTTTTCGACGCTGTTTCCGATCCGATCCGATCCGATCC 128
Qy 625 ACTTGGTTTAAAGATTTCTTACCTGTGGACACAAAGCAACAAATGGTTCGTTAAAGCAG 684
Db 129 ACTTGGTTTAAAGATTTCTTACCTGTGGACACAAAGCAACAAATGGTTCGTTAAAGCAG 188
Qy 685 TTACGATCAGAAATCTATTGGTGGTATACACAAATAGAGAGGAGCCCTTCAGATTGAGCAGT 744
Db 189 TTACGATCAGACCCCA--GGTGGTACACAAATAGAGAGGAGCCCTTCAGATTGAGCAGT 246
Qy 745 GAAGAGTCTGACCAAGGAAATATCAGTGTGTTGCCACCAACAGCGCGGCACTCGCTAT 804
Db 247 GAAGAGTCTGACCAAGGAAATATCAGTGTGTTGCCACCAACAGCGCGGCACTCGCTAT 306
Qy 805 TCCGCTCTCCCAAATTTATATGTCAGAGAGCTGCGAGAAAGTTTCGCGGTGCCACCAAGA 864
Db TCCGCTCTCCCAAATTTATATGTCAGAGAGCTGCGAGAAAGTTTCGCGGTGCCACCAAGA 1442
```

```
Db 307 TCCGCTCTCCCAAATTTATATGTCAGAG---GTAGGAATTTTCGCGGTGCCACCAAGA 362
Qy 865 TTCTCTATCCCAACCACTAATCATGAATCATGCCAGGCGGAAGCGTTTATATATCACCTGT 924
Db 363 TTCTCTATCCCAACCACTAATCATGAATCATGCCAGGCGGAAGCGTTTATATATCACCTGT 422
Qy 925 GTGGCGTGGGGTCCCAATGCTTATGTAAGTGGATGTTGGGGGAGAAAGATCTGACA 984
Db 423 GTGGCGTGGGGTCCCAATGCTTATGTAAGTGGATGTTGGGGGAGAAAGATCTGACA 482
Qy 985 CTTGAAGATGATATGCCAATAGGAAGAAATGTCTAGAACTGAATGATGTAAGACAGTCA 1044
Db 483 CTTGAAGATGATATGCCAATAGGAAGAAATGTCTAGAACTGAATGATGTAAGACAGTCA 542
Qy 1045 GCAAAATTACACCTGTGTGCTATGTCAAACCTGGGTGTGTAAGCAATAGCACAGATC 1104
Db 543 GCAAAATTACACCTGTGTGCTATGTCAAACCTGGGTGTGTAAGCAATAGCACAGATC 602
Qy 1105 ACTGTCAAAGCCTTACCCAAACCTCCAGGAACCTCTGTAGTAGCCAGAGACAGCTACA 1164
Db 603 ACTGTCAAAGCCTTACCCAAACCTCCAGGAACCTCTGTAGTAGCCAGAGACAGCTACA 662
Qy 1165 AGCATCACACTGACGTGGGACTCTGGGAACCTCTGAGCCTGTTTCTTATTACATAATTCAG 1224
Db 663 AGCATCACACTGACGTGGGACTCTGGGAACCTCTGAGCCTGTTTCTTATTACATAATTCAG 722
Qy 1225 CATAAACCTTAAACCTCTGAGGAACCTTACAAAGAAATTTGATGGGGTGGGACACACGC 1284
Db 723 CATAAACCTTAAACCTCTGAGGAACCTTACAAAGAAATTTGATGGGGTGGGACACACGC 782
Qy 1285 TACAGTGTCTGCTGAGCTAAGTCCCTACTCGGATTTATGAATTCAGGGTGTGTTGCTCAAT 1344
Db 783 TACAGTGTCTGCTGAGCTAAGTCCCTACTCGGATTTATGAATTCAGGGTGTGTTGCTCAAT 842
Qy 1345 AACATTTGGCGGGGGCTCCACGCAACCTGTCTAACCAAACTCAGAGCAAGCACCA 1404
Db 843 AACATTTGGCGGGGGCTCCACGCAACCTGTCTAACCAAACTCAGAGCAAGCACCA 902
Qy 1405 TCCAGTGTCCCGAGGAGTGTCCAGGACCAAGATTTGAGTTTCGACCACTTTTGGTACAG 1464
Db 903 TCCAGTGTCCCGAGGAGTGTCCAGGACCAAGATTTGAGTTTCGACCACTTTTGGTACAG 962
Qy 1465 TCGAAGGAACCTGAAGAGCCAAATGGACAGATCAAGGATATAGAGTTTATTATACAAATG 1524
Db 963 TCGAAGGAACCTGAAGAGCCAAATGGACAGATCAAGGATATAGAGTTTATTATACAAATG 1022
Qy 1525 GATCCCACTCAACATGTCAAACAACTGGATGAAACACAAATGTAGTGACAGCCAAATCACT 1584
Db 1023 GATCCCACTCAACATGTCAAACAACTGGATGAAACACAAATGTAGTGACAGCCAAATCACT 1082
Qy 1585 ACTATTGGCAACTTAGTGGCCCAAGAAACATATTCTGCAAGTCCGCTTTTACCTCA 1644
Db 1083 ACTATTGGCAACTTAGTGGCCCAAGAAACATATTCTGCAAGTCCGCTTTTACCTCA 1142
Qy 1645 ATTGGAGATGGTCCCTTTTCAAGTGACATCAAGTCACTCAGACAGGAGTACCAGGG 1704
Db 1143 ATTGGAGATGGTCCCTTTTCAAGTGACATCAAGTCACTCAGACAGGAGTACCAGGG 1202
Qy 1705 CAGCCACTAAACTTTCAAAGCAGAACCTGAGTCTGTAACAAAGATTTTCTCTTGACA 1764
Db 1203 CAGCCACTAAACTTTCAAAGCAGAACCTGAGTCTGTAACAAAGATTTTCTCTTGACA 1262
Qy 1765 CTTCCAGTTCAGATACCAATTTGCCAATGATGAACTGTCTACAAAGATGGGAGCATGGA 1824
Db 1263 CTTCCAGTTCAGATACCAATTTGCCAATGATGAACTGTCTACAAAGATGGGAGCATGGA 1322
Qy 1825 GAGGAGCAAGAAATTAACATTTAGCCAGGAGACATATATAGGCTGCAAGGACTGAAACCA 1884
Db 1323 GAGGAGCAAGAAATTAACATTTAGCCAGGAGACATATATAGGCTGCAAGGACTGAAACCA 1382
Qy 1885 AACAGCTTATATCTATTTCCTGCTGGCTGACAGCTCCCTCAAGGCTGGGTGCTTCTACT 1944
Db 1383 AACAGCTTATATCTATTTCCTGCTGGCTGACAGCTCCCTCAAGGCTGGGTGCTTCTACT 1442
```

QY 1945 GCAGAAATATCAGCTAGAACCAATGCAATCAAGCCGTCAGCTCCTCCTCAAGACATTAAGT 2004
DB 1443 GCAGAAATATCAGCTAGAACCAATGCAATCAAGCCGTCAGCTCCTCCTCAAGACATTAAGT 1502
QY 2005 TGCACAGCCCAAGTTTCCACTAGTATTTTGGTAAGTTGGCAACCTCCACAGCTGGAATAA 2064
DB 1503 TGCACAGCCCAAGTTTCCACTAGTATTTTGGTAAGTTGGCAACCTCCACAGCTGGAATAA 1562
QY 2065 CAGAAATGGCAATATCACTGAATATCTCAATCAAGTACACTGCAAGTGGATGGGGAAGATGAC 2124
DB 1563 CAGAAATGGCAATATCACTGAATATCTCAATCAAGTACACTGCAAGTGGATGGGGAAGATGAC 1622
QY 2125 AAGCCTCAGGAGATTTTGGGAATTTCTTGGGACACTACCAATATACCTTTTGGGAACAGCTG 2184
DB 1623 AAGCCTCAGGAGATTTTGGGAATTTCTTGGGACACTACCAATATACCTTTTGGGAACAGCTG 1682
QY 2185 GAAAAATGGCAATATCACTGAATATCTCAATCAAGTACACTGCAAGTGGATGGGGAAGATGAC 2244
DB 1683 GAAAAATGGCAATATCACTGAATATCTCAATCAAGTACACTGCAAGTGGATGGGGAAGATGAC 1742
QY 2245 GAGAGCTTGTCCGTTGATTCGAAACCAATGAAGATGTTCTAGTGGTCTCTCCTCGCAAA 2304
DB 1743 GAGAGCTTGTCCGTTGATTCGAAACCAATGAAGATGTTCTAGTGGTCTCTCCTCGCAAA 1802
QY 2305 GTCGAGGTAGAGGCTGTCACTCAACATCTGTTAAAGTCTCATGCGGCTCACCCTGCGCC 2364
DB 1803 GTCGAGGTAGAGGCTGTCACTCAACATCTGTTAAAGTCTCATGCGGCTCACCCTGCGCC 1862
QY 2365 AATAAACAGCATGGCCAGATAGAGGATATCAGGTGCAATATGTGAGGATGGAAATGAT 2424
DB 1863 AATAAACAGCATGGCCAGATAGAGGATATCAGGTGCAATATGTGAGGATGGAAATGAT 1922
QY 2425 GAGCCAAAGGCGACGCCATGCTGAAAGATGTCATGCTGGCTGATGACACA--GGATATTT 2484
DB 1923 GAGCCAAAGGCGACGCCATGCTGAAAGATGTCATGCTGGCTGATGACACA--GGATATTT 1980
QY 2485 GATGATACATGAAACATGACATGATCATTTCTGGGCTCCAGCTGAAACCTTCTACTCC 2544
DB 1981 GATGAAACT-----CATTTACATGATCATTTCTGGGCTCCAGCTGAAACCTTCTACTCC 2034
QY 2545 CTCACCGTCACAGCTTACACAAACAGAGATGCTGCTCGCAGCAAGCCCAAACTGGTG 2604
DB 2035 CTCACCGTCACAGCTTACACAAACAGAGATGCTGCTCGCAGCAAGCCCAAACTGGTG 2094
QY 2605 TCCACCACTGGGGCAGTTCCAGGGAAACCTCGGCTTGTGATTAACCACTCAGATGAAT 2664
DB 2095 TCCACCACTGGGGCAGTTCCAGGGAAACCTCGGCTTGTGATTAACCACTCAGATGAAT 2154
QY 2665 ACTGCTCTTATTCAGTGGCACTTCCGCTGGGACACATTTGGACCTTTCAGGGCTACCGT 2724
DB 2155 ACTGCTCTTATTCAGTGGCACTTCCGCTGGGACACATTTGGACCTTTCAGGGCTACCGT 2214
QY 2725 CTAAATTTGGCCGCAAGGATATGAGGCACTTACTCTTGTAGTCTCTGAAAGAAAGAA 2784
DB 2215 CTAAATTTGGCCGCAAGGATATGAGGCACTTACTCTTGTAGTCTCTGAAAGAAAGAA 2274
QY 2785 GATCACTTTACAGCTACAGACATCCAAAGGGAGCATCATAGTCTTCAGGCTCTCAGCC 2844
DB 2275 GATCACTTTACAGCTACAGACATCCAAAGGGAGCATCATAGTCTTCAGGCTCTCAGCC 2334
QY 2845 AGAAACAAAGTGGGCTTTGGGGAGGAGATGTTGAAGGAGATTTCCATTTCCAGAGAGATA 2904
DB 2335 AGAAACAAAGTGGGCTTTGGGGAGGAGATGTTGAAGGAGATTTCCATTTCCAGAGAGATA 2394
QY 2905 CCAACTGATTTCCCTCAAAACCTTCACTCAGAGGCCACCTTCAACCTCCGCTCCAGTTA 2964
DB 2395 CCAACTGATTTCCCTCAAAACCTTCACTCAGAGGCCACCTTCAACCTCCGCTCCAGTTA 2454
QY 2965 TCTTGGCAACCACTGCTCGGAGAGAGAAATGGCAATATCACCAGTATACCTTCTT 3024
DB 2455 TCTTGGCAACCACTGCTCGGAGAGAGAAATGGCAATATCACCAGTATACCTTCTT 2514

QY 3025 TATAGGAGATCAACATCCCTCTCTCCGATGGAGGAGCTTATTTGTTCCAGCTGACACC 3084
DB 2515 TATAGGAGATCAACATCCCTCTCTCCGATGGAGGAGCTTATTTGTTCCAGCTGACACC 2574
QY 3085 ACTATGACACTCACTGGCTTAAACCAGATACACATACGATGTAAAAAGTACGCTCAT 3144
DB 2575 ACTATGACACTCACTGGCTTAAACCAGATACACATACGATGTAAAAAGTACGCTCAT 2634
QY 3145 ACAGAGAAAGGCGCGGCCCATATAGTCCAGTGTCCAGTGTCCAGGACACTGCTGTGAT 3204
DB 2635 ACAGAGAAAGGCGCGGCCCATATAGTCCAGTGTCCAGTGTCCAGGACACTGCTGTGAT 2694
QY 3205 CAAGTGTTCGAAATAATTTTCATGTCAAGCAGTAAATGAAGACTTCCGTTGCTGTCT 3264
DB 2695 CAAGTGTTCGAAATAATTTTCATGTCAAGCAGTAAATGAAGACTTCCGTTGCTGTCT 2754
QY 3265 TGGGAGATTCAGAGAAATTAATACTCCGCCATGCTTTCAAAAATCTTTATGATGATGG 3324
DB 2755 TGGGAGATTCAGAGAAATTAATACTCCGCCATGCTTTCAAAAATCTTTATGATGATGG 2814
QY 3325 AAAATGTTAGAAAGTGGATGGCGGACCAACAG--AAGTTAAATGTCAACCTGAAGCC 3383
DB 2815 AAAATGTTAGAAAGTGGATGGCGGACCAACAGAAAGTTAAATGTCAACCTGAAGCC 2874
QY 3384 TGAGAAATCATATTTTCATTTGCTGACAAATCGTGGAAACAGTGTGCTGGCTGACGA 3443
DB 2875 TGAGAAATCATATTTTCATTTGCTGACAAATCGTGGAAACAGTGTGCTGGCTGACGA 2934
QY 3444 CAGGCTCAGGCAAGACTGCAACAGATGTATTAAGTACCAAGCTGCTTTTATTGGGAA 3503
DB 2935 CAGGCTCAGGCAAGACTGCAACAGATGTATTAAGTACCAAGCTGCTTTTATTGGGAA 2994
QY 3504 GACCAACTGGATGGCATGATTTACTGTGCAACTGCTGAACTGCTGCAAAATGAGAAAT 3563
DB 2995 GACCAACTGGATGGCATGATTTACTGTGCAACTGCTGAACTGCTGCAAAATGAGAAAT 3054
QY 3564 AAAAGGTTACTACATATATTTGCTGCTTTGAAAGAAATCTCGCGGAAATTTTCAAGCC 3623
DB 3055 AAAAGGTTACTACATATATTTGCTGCTTTGAAAGAAATCTCGCGGAAATTTTCAAGCC 3114
QY 3624 ATGGAGAGTCCAGATCAAAATGGAATTTAGATGAGTCTGCTTAAAGGAGATATCTAGGAAGCG 3683
DB 3115 ATGGAGAGTCCAGATCAAAATGGAATTTAGATGAGTCTGCTTAAAGGAGATATCTAGGAAGCG 3174
QY 3684 CAGAAGCATCCGTTATGGGAGAGAAAGTTGAATTTAAAGCCATATATTCGCTCACTTTGA 3743
DB 3175 CAGAAGCATCCGTTATGGGAGAGAAAGTTGAATTTAAAGCCATATATTCGCTCACTTTGA 3234
QY 3744 TGTCTTCCCACTGAGTTACCTCGGGGATGACAAGCATTTATGGTGAATTTACAAACAA 3803
DB 3235 TGTCTTCCCACTGAGTTACCTCGGGGATGACAAGCATTTATGGTGAATTTACAAACAA 3294
QY 3804 GCAACTCAAAAGTGGTCAAGAAATATGCTTCTTTGTTAGCAGTAAATGGAAACATGACAGA 3863
DB 3295 GCAACTCAAAAGTGGTCAAGAAATATGCTTCTTTGTTAGCAGTAAATGGAAACATGACAGA 3354
QY 3864 GTCTAAGATGTATGCAACCAAGCCCTTACTCGACCCCGTGGTGTCAATGGATCTGGATCC 3923
DB 3355 GTCTAAGATGTATGCAACCAAGCCCTTACTCGACCCCGTGGTGTCAATGGATCTGGATCC 3414
QY 3924 GCAGCCAAATCAGGATCAAGAAAGGCTTGAATCTGGGTTGTAGGTTCTGCTTGGCAGT 3983
DB 3415 GCAGCCAAATCAGGATCAAGAAAGGCTTGAATCTGGGTTGTAGGTTCTGCTTGGCAGT 3474
QY 3984 GGTCTTTTATCATCTGCAATTTGCTATTGCTATTCTTTTATAAAGGAGAGAGGCGAGAGTC 4043
DB 3475 GGTCTTTTATCATCTGCAATTTGCTATTGCTATTCTTTTATAAAGGAGAGAGGCGAGAGTC 3534
QY 4044 CGACTCTAGAAAAAGCAGCATACCGAAACAATAAGGAGATCCCTTTCAACCAACCAACAGA 4103
DB 3535 CGACTCTAGAAAAAGCAGCATACCGAAACAATAAGGAGATCCCTTTCAACCAACCAACAGA 3594
QY 4104 CCCTGTAGAACTGAGGGCGCTTAACTTTCAAAACACCGGGTATGGCTAGCCATCTCCAT 4163

3595	Db	CCCTGTAGAACTCGAGGCGCCTTAACTTTCAACAACCGGGTATGGCTAGCCATCCTCCAAT	3654
4164	Qy	ACCCATCTTCGAACTTCGACGACCACATTTGAAAGATTTGAAAGCAAAATGACAACATTGAAGTT	4223
3655	Db	ACCCATCTTCGAACTTCGACGACCACATTTGAAAGATTTGAAAGCAAAATGACAACATTGAAGTT	3714
4224	Qy	TTCCAGGAATATGAGTCAATTGACCCCTGGCCAGCAGTTTCACTTTGGGAAACATTCAAACTT	4283
3715	Db	TTCCAGGAATATGAGTCAATTGACCCCTGGCCAGCAGTTTCACTTTGGGAAACATTCAAACTT	3774
4284	Qy	GGAGTTAAACAAACCAAAGAAATAGATACGGGAATGTTAATCGCATATGATCATTTCCCGGGT	4343
3775	Db	GGAGTTAAACAAACCAAAGAAATAGATACGGGAATGTTAATCGCATATGATCATTTCCCGGGT	3834
4344	Qy	TCTCCTATCAGCTATAGAAAGGATCCAGGAGTGAATGTGAAATGCCCACTACATAGAGA	4403
3835	Db	TCTCCTATCAGCTATAGAAAGGATCCAGGAGTGAATGTGAAATGCCCACTACATAGAGA	3894
4404	Qy	TGGGTATAGGAAGCAAAATGCCTATATTTGCAACACAGGGATCTCTCCCGGAAACATTTGG	4463
3895	Db	TGGGTATAGGAAGCAAAATGCCTATATTTGCAACACAGGGATCTCTCCCGGAAACATTTGG	3954
4464	Qy	GGACTTTTGGAGAAATGATATATGGGAACAAACGGAGTGCACAGTTGTGTATGATGACAAACT	4523
3955	Db	GGACTTTTGGAGAAATGATATATGGGAACAAACGGAGTGCACAGTTGTGTATGATGACAAACT	4014
4524	Qy	AGAAGAAAGATCAAGGCTGAAGTGTGACCCAGTATTTGGCCTAGCAGAGGCAACAGAACCCA	4583
4015	Db	AGAAGAAAGATCAAGGCTGAAGTGTGACCCAGTATTTGGCCTAGCAGAGGCAACAGAACCCA	4074
4584	Qy	CGGACTCGTTCAAGTAAACCGCTGCTTGATATCTGTGGAGCTGGCCACATATTTGTGTTCGAAC	4643
4075	Db	CGGACTCGTTCAAGTAAACCGCTGCTTGATATCTGTGGAGCTGGCCACATATTTGTGTTCGAAC	4134
4644	Qy	ATTTGCACTTTTACAAGAAATGGTTCAGTGAAGAGAGAAAGTGAACAATTCAGATTTCAC	4703
4135	Db	ATTTGCACTTTTACAAGAAATGGTTCAGTGAAGAGAGAAAGTGAACAATTCAGATTTCAC	4194
4704	Qy	CGCCTGGCTGATCATCTGTGTTTCAGAACACCCCTACACCTTTTCTAGCTTCTTACGTAG	4763
4195	Db	CGCCTGGCTGATCATCTGTGTTTCAGAACACCCCTACACCTTTTCTAGCTTCTTACGTAG	4254
4764	Qy	AGTCAAAACCTGTAAACCCCTCCGATGCTGGTCCGATGGTTGTGCACCTGCAGTGGCGAGT	4823
4255	Db	AGTCAAAACCTGTAAACCCCTCCGATGCTGGTCCGATGGTTGTGCACCTGCAGTGGCGAGT	4314
4824	Qy	TGGCCGGAATGGTTGCTTCATCGTATAGATGCCATGTTAGAAAGAAATAAGCATGAAAA	4883
4315	Db	TGGCCGGAATGGTTGCTTCATCGTATAGATGCCATGTTAGAAAGAAATAAGCATGAAAA	4374
4884	Qy	AACCTGTAGATATTTATGGCCATGTAACTTTAATCAGAGCCGAGGAACTATATGTTTCA	4943
4375	Db	AACCTGTAGATATTTATGGCCATGTAACTTTAATCAGAGCCGAGGAACTATATGTTTCA	4434
4944	Qy	AACAGAGACCAATACATCTTTATTCATGATGCACTGTTAGAAGCAGTGAATCTGTGAAA	5003
4435	Db	AACAGAGACCAATACATCTTTATTCATGATGCACTGTTAGAAGCAGTGAATCTGTGAAA	4494
5004	Qy	TACCGAAGTGCAGCTAGAACTTTGTATGCTTCATATTGAGAGCTGACACAAATAGAAAC	5063
4495	Db	TACCGAAGTGCAGCTAGAACTTTGTATGCTTCATATTGAGAGCTGACACAAATAGAAAC	4554
5064	Qy	GGGAGAGATGTGACAGGAATGGAGCTCGAAATTTAAGGCTTACGACAGCTCAAAAGCTCA	5123
4555	Db	GGGAGAGATGTGACAGGAATGGAGCTCGAAATTTAAGGCTTACGACAGCTCAAAAGCTCA	4614
5124	Qy	CACCTCAAGGTTTATCAGTGCCTAATCTTTCATGTAAATAATTTCAAAATCGCCTTGTATA	5183
4615	Db	CACCTCAAGGTTTATCAGTGCCTAATCTTTCATGTAAATAATTTCAAAATCGCCTTGTATA	4674
5184	Qy	TATTATGTCATATGAATCCAAAGGGTATGCTCGAGCTTATCGTGGAGTAGAAGGATC	5243

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 5, 2005, 22:23:57 ; Search time 923 Seconds
(without alignments)
11102.926 Million cell updates/sec

Title: US-10-772-636-63
Perfect score: 6263
Sequence: 1 gctaactcaaggagagcgc.....aagtgaattttaaaaaagttt 6263

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6263	100.0	6263	4	US-09-023-655-1473
2	2473	39.5	6363	4	Sequence 1473, Ap
3	2398	38.3	6000	1	Sequence 879, App
4	2398	38.3	6000	2	Sequence 6, Appli
5	2398	38.3	6000	3	Sequence 6, Appli
6	2398	38.3	6000	5	Sequence 6, Appli
7	2357.4	37.6	7702	4	Sequence 1336, Ap
8	2357.4	37.6	7702	4	Sequence 1336, Ap
9	1345.8	21.5	4078	4	Sequence 1132, Ap
10	1300	20.8	5690	2	Sequence 2, Appli
11	1300	20.8	5690	2	Sequence 2, Appli
12	1095.8	17.5	3467	4	Sequence 1, Appli
13	538	8.6	1029	4	Sequence 23, Appli
14	538	8.6	1029	4	Sequence 24, Appli
15	442	7.1	2872	1	Sequence 4, Appli
16	442	7.1	2872	4	Sequence 4, Appli
17	438.8	7.0	2409	1	Sequence 2, Appli
18	438.8	7.0	2409	4	Sequence 2, Appli
19	438.8	7.0	2537	4	Sequence 1485, Ap
20	438.8	7.0	2565	4	Sequence 1311, Ap
21	438.8	7.0	3615	4	Sequence 1485, Ap
22	376.8	6.0	1689	4	Sequence 1254, Ap
23	348	5.6	4651	2	Sequence 4915, Ap
24	348	5.6	4651	2	Sequence 4, Appli
25	348	5.6	4651	2	Sequence 4, Appli
26	338.4	5.4	4374	2	Sequence 945, App
27	338.4	5.4	4374	2	Sequence 3, Appli

28 321.4 5.1 4078 4 US-09-016-434-1134 Sequence 1134, Ap
29 317.2 5.1 3228 4 US-09-949-016-2437 Sequence 2437, Ap
30 317.2 5.1 4998 4 US-09-023-655-1299 Sequence 1299, Ap
31 270 4.3 5769 1 US-08-652-971-1 Sequence 1, Appli
32 270 4.3 5769 2 US-08-991-258A-1 Sequence 1, Appli
33 270 4.3 5769 2 US-08-769-399-1 Sequence 1, Appli
34 270 4.3 5769 3 US-08-991-953A-1 Sequence 1, Appli
35 267.4 4.3 2422 4 US-09-949-016-1287 Sequence 1287, Ap
36 265 4.2 7941 3 US-09-816-703A-1 Sequence 1, Appli
37 260.2 4.2 6924 1 US-08-015-973-2 Sequence 2, Appli
38 260.2 4.2 6924 2 US-08-448-164-2 Sequence 2, Appli
39 259.6 4.1 3995 4 US-09-949-016-2026 Sequence 2026, Ap
40 259.6 4.1 4338 1 US-08-015-986A-1 Sequence 1, Appli
41 259.6 4.1 4338 2 US-08-446-363-1 Sequence 1, Appli
42 257.8 4.1 4389 4 US-09-023-655-991 Sequence 991, App
43 256.6 4.1 5581 4 US-09-023-655-966 Sequence 966, App
44 220.2 3.5 4798 4 US-09-620-312D-298 Sequence 298, App
45 218.8 3.5 2520 4 US-09-949-016-4698 Sequence 4698, Ap

ALIGNMENTS

RESULT 1
US-09-023-655-1473
; Sequence 1473, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1473:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6263 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g755652
US-09-023-655-1473

Query Match 100.0% ; Score 6263 ; DB 4 ; Length 6263 ;

Best Local Similarity 100.0%; Pred. No. 0; Matches 6263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	GCTAACTCAAGGGAGAGCGTCTGTGGTGAACACCCGCTGGGATCTAAAGAACAAAGCTCTGAAG	60
Db	1	GCTAACTCAAGGGAGAGCGTCTGTGGTGAACACCCGCTGGGATCTAAAGAACAAAGCTCTGAAG	60
Qy	61	TGTTCCAGCTGAATTTTCAGATCGGACAGACTCGCTCGGCTCCGGAGGACGATGATCCA	120
Db	61	TGTTCCAGCTGAATTTTCAGATCGGACAGACTCGCTCGGCTCCGGAGGACGATGATCCA	120
Qy	121	AGCTGCTCGGCGACGCTGCTCCAAAGCTGCGAGGATGTCAGCGTAGCGAGCGCTCTGCTG	180
Db	121	AGCTGCTCGGCGACGCTGCTCCAAAGCTGCGAGGATGTCAGCGTAGCGAGCGCTCTGCTG	180
Qy	181	CTGCTCTCTCACTTTCTCTCCGACACGATGCTGAGACACCTCCAAAGGTTTACACGAACA	240
Db	181	CTGCTCTCTCACTTTCTCTCCGACACGATGCTGAGACACCTCCAAAGGTTTACACGAACA	240
Qy	241	CCGTTGATACAGACGGGCTCTGGCGGAGTTGCTCTTTCATCTGCGAAGCTACGGGA	300
Db	241	CCGTTGATACAGACGGGCTCTGGCGGAGTTGCTCTTTCATCTGCGAAGCTACGGGA	300
Qy	301	GACCCAAAGCTTAAATTTCTCGAAACAAAGGAGGAAAGTCAAGATTCAGAGATTT	360
Db	301	GACCCAAAGCTTAAATTTCTCGAAACAAAGGAGGAAAGTCAAGATTCAGAGATTT	360
Qy	361	GAGTAAATAGAGTTTGACGATGGTCTGGATCAGTTCTCAGAAATACACCCCTTACGGACT	420
Db	361	GAGTAAATAGAGTTTGACGATGGTCTGGATCAGTTCTCAGAAATACACCCCTTACGGACT	420
Qy	421	CCGAGGATAGAGCCATTTATGAATGTGGCCCTCAATATGTGGAGAAATTAAGTGTA	480
Db	421	CCGAGGATAGAGCCATTTATGAATGTGGCCCTCAATATGTGGAGAAATTAAGTGTA	480
Qy	481	TCCACACAGCTCAGATTTTGCGGAGAGATCAAAATCCAGGGCTTCCCTTACCATTTGAC	540
Db	481	TCCACACAGCTCAGATTTTGCGGAGAGATCAAAATCCAGGGCTTCCCTTACCATTTGAC	540
Qy	541	ATGGGCCCAAGTTGAAAGTGTGTGAGCGTACTCGACGGCCACCATGCTTTGTGACGCC	600
Db	541	ATGGGCCCAAGTTGAAAGTGTGTGAGCGTACTCGACGGCCACCATGCTTTGTGACGCC	600
Qy	601	AGTGGTAATCCGGATCCAGAAATCACTTGGTTTAAAGATTTCTTACCTGTGGACACAAGC	660
Db	601	AGTGGTAATCCGGATCCAGAAATCACTTGGTTTAAAGATTTCTTACCTGTGGACACAAGC	660
Qy	661	AACAAATGTCGTATTAAAGCAGTTACGATCAGAAATCTATTGGTGTACACCAATTAAGA	720
Db	661	AACAAATGTCGTATTAAAGCAGTTACGATCAGAAATCTATTGGTGTACACCAATTAAGA	720
Qy	721	GGAGCCCTTCAGATTGAGCAGAGTGAAGAGTCTGACCAAGGAAATATGAGTGTGTGCC	780
Db	721	GGAGCCCTTCAGATTGAGCAGAGTGAAGAGTCTGACCAAGGAAATATGAGTGTGTGCC	780
Qy	781	ACCAACAGCGGGGACCTCGCTATTTCGCTCTCTGCGCAATTTATATGTACAGAGCTCGGA	840
Db	781	ACCAACAGCGGGGACCTCGCTATTTCGCTCTCTGCGCAATTTATATGTACAGAGCTCGGA	840
Qy	841	GAGTGTGCGCTGTGCCACCAAGATTTCTATCCACCCTCAATCATGAATCATGCGCA	900
Db	841	GAGTGTGCGCTGTGCCACCAAGATTTCTATCCACCCTCAATCATGAATCATGCGCA	900
Qy	901	GGCGGAAGCGTTAATATCACCTGTGTGCGCGGTGAGGATGATGCCAATGGAAGAAATGTGCTA	960
Db	901	GGCGGAAGCGTTAATATCACCTGTGTGCGCGGTGAGGATGATGCCAATGGAAGAAATGTGCTA	960
Qy	961	ATGTTGGGGCAGAGATCTGACACCTGGAAGATGATGCCAATGGAAGAAATGTGCTA	1020
Db	961	ATGTTGGGGCAGAGATCTGACACCTGGAAGATGATGCCAATGGAAGAAATGTGCTA	1020
Qy	1021	GAACTGAATGATGTAAGACAGTCAAGAAATTAACCTGTGTGTATGTCAACACTGGGT	1080
Db	1021	GAACTGAATGATGTAAGACAGTCAAGAAATTAACCTGTGTGTATGTCAACACTGGGT	1080
Qy	1081	GTCAATTGAAGCAATAGCAGATCACTGTCAAAAGCCTTACCAAAACCTCCAGGAACTCCT	1140
Db	1081	GTCAATTGAAGCAATAGCAGATCACTGTCAAAAGCCTTACCAAAACCTCCAGGAACTCCT	1140
Qy	1141	GTAGTGACCGAGACAGCTCAAGCATCACTGACGTGGGACTCTGGGAAACCTCTGAG	1200
Db	1141	GTAGTGACCGAGACAGCTCAAGCATCACTGACGTGGGACTCTGGGAAACCTCTGAG	1200
Qy	1201	CCTGTTTCTTATTACATAATTCAGCATAAACCTTAAAAAACCTCTGAGGAACTTTACAAAGAA	1260
Db	1201	CCTGTTTCTTATTACATAATTCAGCATAAACCTTAAAAAACCTCTGAGGAACTTTACAAAGAA	1260
Qy	1261	ATTGATGGGGTGGCGACCAACACGCTACAGTGTGCTGGACTAAGTCTCTACTCTCGGATAT	1320
Db	1261	ATTGATGGGGTGGCGACCAACACGCTACAGTGTGCTGGACTAAGTCTCTACTCTCGGATAT	1320
Qy	1321	GAATTCAGGGTGTGCTGCTCAATTAACATTTGGGGGGGGCTCCACAGGAACTTGTGCTA	1380
Db	1321	GAATTCAGGGTGTGCTGCTCAATTAACATTTGGGGGGGGCTCCACAGGAACTTGTGCTA	1380
Qy	1381	ACAAACCTTCAGAGCAAGCAACCATCCAGTGTGCTGGAGGATGTCCAGGCAAGATGTTG	1440
Db	1381	ACAAACCTTCAGAGCAAGCAACCATCCAGTGTGCTGGAGGATGTCCAGGCAAGATGTTG	1440
Qy	1441	AGTTGACACCAATTTTGTGTACAGTGGAGGAAACCTGAAAGGCCAAATGGACAGATCCAA	1500
Db	1441	AGTTGACACCAATTTTGTGTACAGTGGAGGAAACCTGAAAGGCCAAATGGACAGATCCAA	1500
Qy	1501	GGATATAGAGTTTATTATACAATGGATCCCACTCAACATGTCAACACTGGATGAAACAC	1560
Db	1501	GGATATAGAGTTTATTATACAATGGATCCCACTCAACATGTCAACACTGGATGAAACAC	1560
Qy	1561	AATGTAGCTCAGACCAAAATCACTATTATTTGGCAACTTAGTGCCCCAGAAAACATATTTCT	1620
Db	1561	AATGTAGCTCAGACCAAAATCACTATTATTTGGCAACTTAGTGCCCCAGAAAACATATTTCT	1620
Qy	1621	GTCAAGTCTCTGGCTTTTACTCAATTTGGAGATGGTCCCTTTCAAGTGACATACAAGTC	1680
Db	1621	GTCAAGTCTCTGGCTTTTACTCAATTTGGAGATGGTCCCTTTCAAGTGACATACAAGTC	1680
Qy	1681	ATCACTCAGACAGAGTACAGGCGCAGCTTAACTTCAAAAGCAGAACCTTGAGCTGTGAA	1740
Db	1681	ATCACTCAGACAGAGTACAGGCGCAGCTTAACTTCAAAAGCAGAACCTTGAGCTGTGAA	1740
Qy	1741	ACAAGTATTTTGTCTCTTGGACACCTCCAGTTCAGATACCAATTTGCCAATATGAAGCTG	1800
Db	1741	ACAAGTATTTTGTCTCTTGGACACCTCCAGTTCAGATACCAATTTGCCAATATGAAGCTG	1800
Qy	1801	GTCTACAAAGATGGGAGCATGGAGGAGCAACGAATTTACATTGAGCGGAGACATCA	1860
Db	1801	GTCTACAAAGATGGGAGCATGGAGGAGCAACGAATTTACATTGAGCGGAGACATCA	1860
Qy	1861	TATAGGCTGCAAGSACTGAAACCAACAGCTTACTATTTCCTCTGGCTGCAAGCTCC	1920
Db	1861	TATAGGCTGCAAGSACTGAAACCAACAGCTTACTATTTCCTCTGGCTGCAAGCTCC	1920
Qy	1921	CCTCAAGGCTTGGGTCTCTTACTGCAAGAAATATCAGTAGAACCATGCAAGTCAAGGCG	1980
Db	1921	CCTCAAGGCTTGGGTCTCTTACTGCAAGAAATATCAGTAGAACCATGCAAGTCAAGGCG	1980
Qy	1981	TCAGTCTCTCTCAAGACATTTAGTTGCACAGCCCAAGTTCCTACTAGTATTTTGGTAAGT	2040
Db	1981	TCAGTCTCTCTCAAGACATTTAGTTGCACAGCCCAAGTTCCTACTAGTATTTTGGTAAGT	2040
Qy	2041	TGCAACCTCCACAGTGGAAACAGAAATGGCAATTAATCACTGAATACTCATCAAGTAC	2100
Db	2041	TGCAACCTCCACAGTGGAAACAGAAATGGCAATTAATCACTGAATACTCATCAAGTAC	2100
Qy	2101	ACTGCAAGTGGATGGGAGATGACAGCTTCAAGCTTGGGAAATTTTGGGAAATTTCTTGGGAACT	2160
Db	2101	ACTGCAAGTGGATGGGAGATGACAGCTTCAAGCTTGGGAAATTTTGGGAAATTTCTTGGGAACT	2160

QY 2161 ACCAAATACCTTTTGGAAACAGCTGGAAATGGAATGGAATACCGGATCACTGTGACAGCC 2220
DB |||||
DB 2161 ACCAAATACCTTTTGGAAACAGCTGGAAATGGAATGGAATACCGGATCACTGTGACAGCC 2220
QY 2221 CATACAGATGTCGGCCCTCGCCCTCGAGAGCTTGTCCGTGTTGATTCGAACCAATCAAGAT 2280
DB |||||
DB 2221 CATACAGATGTCGGCCCTCGCCCTCGAGAGCTTGTCCGTGTTGATTCGAACCAATCAAGAT 2280
QY 2281 GTTCTAGTGGTCTCTCTCGCAAAGTCGAGGTAGAGGCTGTCAACTCAACATCTGTAA 2340
DB |||||
DB 2281 GTTCTAGTGGTCTCTCTCGCAAAGTCGAGGTAGAGGCTGTCAACTCAACATCTGTAA 2340
QY 2341 GTCTCATGGCGCTCAACCGTGGCCCAATAAACAGCATGCGCAGATGAAGGATATCAGGTG 2400
DB |||||
DB 2341 GTCTCATGGCGCTCAACCGTGGCCCAATAAACAGCATGCGCAGATGAAGGATATCAGGTG 2400
QY 2401 CATATATGTAGGATGGAATAATGTCAGCCCAAGGCCACGCCCATCTGAAAGATGTCATG 2460
DB |||||
DB 2401 CATATATGTAGGATGGAATAATGTCAGCCCAAGGCCACGCCCATCTGAAAGATGTCATG 2460
QY 2461 CTGGCTGATGCAAGTGGGAATTTGATGATACTACTGAACATGACATGATCAATTTCTGGG 2520
DB |||||
DB 2461 CTGGCTGATGCAAGTGGGAATTTGATGATACTACTGAACATGACATGATCAATTTCTGGG 2520
QY 2521 CTCAGCCTGAAACTTCTCTACTCCCTCAACCGTCAAGCCTTACACAAACAAAGGAGATGGT 2580
DB |||||
DB 2521 CTCAGCCTGAAACTTCTCTACTCCCTCAACCGTCAAGCCTTACACAAACAAAGGAGATGGT 2580
QY 2581 GCTCCACAGCCCAACCTGCTGTCACCACTGGGGCAGTTCACAGGGAACCTCGGCTT 2640
DB |||||
DB 2581 GCTCCACAGCCCAACCTGCTGTCACCACTGGGGCAGTTCACAGGGAACCTCGGCTT 2640
QY 2641 GTGATTAACCACTCAGATGAATCTGCTCTTAATTCAGTGGCACCTCCCGTGGACACA 2700
DB |||||
DB 2641 GTGATTAACCACTCAGATGAATCTGCTCTTAATTCAGTGGCACCTCCCGTGGACACA 2700
QY 2701 TTTGGACCTTTCAGGGCTACCGTCTAAATTTGGCCGCAAGGATATGGAGCCACTTACT 2760
DB |||||
DB 2701 TTTGGACCTTTCAGGGCTACCGTCTAAATTTGGCCGCAAGGATATGGAGCCACTTACT 2760
QY 2761 ACTCTTGAGTTCTCTGAAAGAGACATCACTTTTACAGCTACAGACATCCACAAAGGAGCA 2820
DB |||||
DB 2761 ACTCTTGAGTTCTCTGAAAGAGACATCACTTTTACAGCTACAGACATCCACAAAGGAGCA 2820
QY 2821 TCATACGTCTTCAGGCTCTCAGCCAGAAACAAAGTGGGCTTTGGGGAGGAGATGGTGAAG 2880
DB |||||
DB 2821 TCATACGTCTTCAGGCTCTCAGCCAGAAACAAAGTGGGCTTTGGGGAGGAGATGGTGAAG 2880
QY 2881 GAGATTTCCATTCAGAAAGATACCAACTGGATTCCTCAAAACCTTCACTCAGAAAGGC 2940
DB |||||
DB 2881 GAGATTTCCATTCAGAAAGATACCAACTGGATTCCTCAAAACCTTCACTCAGAAAGGC 2940
QY 2941 ACCACTTCAACCTCGCTCCAGTTACTTGGCAACCACTGCTGGCAGAGAGAAATGGC 3000
DB |||||
DB 2941 ACCACTTCAACCTCGCTCCAGTTACTTGGCAACCACTGCTGGCAGAGAGAAATGGC 3000
QY 3001 ATTATCAACAAATPACCTTTCTTTATAGGATATCAACATCCCGCTTCTCCCGATGGAG 3060
DB |||||
DB 3001 ATTATCAACAAATPACCTTTCTTTATAGGATATCAACATCCCGCTTCTCCCGATGGAG 3060
QY 3061 CAGCTTATGTTTCAGCTGACACCACTATGACATCTGCTGGGTTAAACCAAGATACCA 3120
DB |||||
DB 3061 CAGCTTATGTTTCAGCTGACACCACTATGACATCTGCTGGGTTAAACCAAGATACCA 3120
QY 3121 TACGATTAAGATACGTGCTCATACGAGCAAGAGGCCCGGCCATATAGTCCCAAGTGC 3180
DB |||||
DB 3121 TACGATTAAGATACGTGCTCATACGAGCAAGAGGCCCGGCCATATAGTCCCAAGTGC 3180
QY 3181 CAGTTTCAGGACACTGCTGTCAGTCAAGTGTTCGAAAAATTTTTCATGTCAAAAGCAGTA 3240
DB |||||
DB 3181 CAGTTTCAGGACACTGCTGTCAGTCAAGTGTTCGAAAAATTTTTCATGTCAAAAGCAGTA 3240

QY 3241 ATGAAGACTTCCGTGTTGCTCTCTTGGGAGATTCAGAGAAATTAATACTCCGCCATGCT 3300
DB |||||
DB 3241 ATGAAGACTTCCGTGTTGCTCTCTTGGGAGATTCAGAGAAATTAATACTCCGCCATGCT 3300
QY 3301 TTCAAAATTCCTTTATGATGATGGGAAATGCTAGAAAGTGGATGGCCGAGCCACACAG 3360
DB |||||
DB 3301 TTCAAAATTCCTTTATGATGATGGGAAATGCTAGAAAGTGGATGGCCGAGCCACACAG 3360
QY 3361 AAGTTAATTTGCAACCTGAGAAATCATATATTCATTTGTGCTGACAAATCGTGA 3420
DB |||||
DB 3361 AAGTTAATTTGCAACCTGAGAAATCATATATTCATTTGTGCTGACAAATCGTGA 3420
QY 3421 AACAGTCTGTGGCTGTCAGCACAGGTCACGGCAAAAGACTGACACAGATGATTAAGT 3480
DB |||||
DB 3421 AACAGTCTGTGGCTGTCAGCACAGGTCACGGCAAAAGACTGACACAGATGATTAAGT 3480
QY 3481 ACCAAGCTGCTCTTCAATTTGGGAAGACCAACTTGGATGGCATGATTTGTGCAACTGCT 3540
DB |||||
DB 3481 ACCAAGCTGCTCTTCAATTTGGGAAGACCAACTTGGATGGCATGATTTGTGCAACTGCT 3540
QY 3541 GAAGTACTCTGCAATGAGAAATATAAAGGTTTACTACATAAATTTGTGCTTTGAAGAA 3600
DB |||||
DB 3541 GAAGTACTCTGCAATGAGAAATATAAAGGTTTACTACATAAATTTGTGCTTTGAAGAA 3600
QY 3601 TCTCGGGGAAATTTATCAAGCCATGGGAGATCCAGATGAAATGGAATTAAGTGAAGCTG 3660
DB |||||
DB 3601 TCTCGGGGAAATTTATCAAGCCATGGGAGATCCAGATGAAATGGAATTAAGTGAAGCTG 3660
QY 3661 CTTAAGGAGATATCTAGGAAGCGCAGAGCATCCGTTATGGGAGAGATTTGAATTAAG 3720
DB |||||
DB 3661 CTTAAGGAGATATCTAGGAAGCGCAGAGCATCCGTTATGGGAGAGATTTGAATTAAG 3720
QY 3721 CCATATATTTGCGCTCACTTTGATGTCCTTCCCACTGAGTTTCACTGGGGGATGACAAAG 3780
DB |||||
DB 3721 CCATATATTTGCGCTCACTTTGATGTCCTTCCCACTGAGTTTCACTGGGGGATGACAAAG 3780
QY 3781 CATTTATGGTGAATTAACAACAAGCACTCAAAAGTGGTCAAGAAATATGCTTCTTTGTG 3840
DB |||||
DB 3781 CATTTATGGTGAATTAACAACAAGCACTCAAAAGTGGTCAAGAAATATGCTTCTTTGTG 3840
QY 3841 TTACAGTAAATGGAACATGACAGATCTAAGATGATGCAACACGCTTACTCCGACCCC 3900
DB |||||
DB 3841 TTACAGTAAATGGAACATGACAGATCTAAGATGATGCAACACGCTTACTCCGACCCC 3900
QY 3901 GTGGTGTCAATGGATCTGGATCCGACGCAATCAAGGATGAAAGAGGCTTGTATCTGG 3960
DB |||||
DB 3901 GTGGTGTCAATGGATCTGGATCCGACGCAATCAAGGATGAAAGAGGCTTGTATCTGG 3960
QY 3961 GTTGTAGGCTCTGCTTTCAGATGGTCTTTATCATCTGCAATTTGCTATTTCTTCTT 4020
DB |||||
DB 3961 GTTGTAGGCTCTGCTTTCAGATGGTCTTTATCATCTGCAATTTGCTATTTCTTCTT 4020
QY 4021 TATAAAGGAGGCGCAGAGTCCGACTCTAGAAAAAGCAGCATACCGAAACAAATAGGAG 4080
DB |||||
DB 4021 TATAAAGGAGGCGCAGAGTCCGACTCTAGAAAAAGCAGCATACCGAAACAAATAGGAG 4080
QY 4081 ATCCCTTTCACACCAACCAACAGACCTCTGTAAGTCTGAGGCGCTTAACTTTCAAAACCG 4140
DB |||||
DB 4081 ATCCCTTTCACACCAACCAACAGACCTCTGTAAGTCTGAGGCGCTTAACTTTCAAAACCG 4140
QY 4141 GGTATGGCTAGCCATCTCTCAATACCCATCTTGGAACTTGGAGACCACTGAAAGATTTG 4200
DB |||||
DB 4141 GGTATGGCTAGCCATCTCTCAATACCCATCTTGGAACTTGGAGACCACTGAAAGATTTG 4200
QY 4201 AAAGCAATGCAAACTTTGAAGTTTCCAGGAAATATGAGTCAATTTGACCTTGGCCAGCAG 4260
DB |||||
DB 4201 AAAGCAATGCAAACTTTGAAGTTTCCAGGAAATATGAGTCAATTTGACCTTGGCCAGCAG 4260
QY 4261 TTCACTTTGGGAACATTTCAAACTTTGGAGTAAACAAACCAAGAAATAGATACGCGAATGTA 4320
DB |||||
DB 4261 TTCACTTTGGGAACATTTCAAACTTTGGAGTAAACAAACCAAGAAATAGATACGCGAATGTA 4320
QY 4321 ATCGCATATGATCATTTCCCGGGTTCTCTTATCAGCTATAGAGGGATTCACAGGAAGTGAC 4380

|||||
Db 4321 ATCCGATATGATCATTCCTCCGGGTTCTCTCTATCAGCTATAGAAAGGATCCCGAGGAAGTGAC 4380
Qy 4381 TATGTGAATGCCAACTACATAGATGGGTATAGGAAGCAAAATGCTATATTGCAACACAG 4440
Db 4381 TATGTGAATGCCAACTACATAGATGGGTATAGGAAGCAAAATGCTATATTGCAACACAG 4440
Qy 4441 GGATCTCTCCCGGAAACATTTTGGGGAATTTTGGGGAATGATATGGAAGCAACAGGAGTGCC 4500
Db 4441 GGATCTCTCCCGGAAACATTTTGGGGAATTTTGGGGAATGATATGGAAGCAACAGGAGTGCC 4500
Qy 4501 ACAGTTGTCTGATGACAAATACTAGAAAGATCAAGGGTGAAGTGTGACCAAGTATTGG 4560
Db 4501 ACAGTTGTCTGATGACAAATACTAGAAAGATCAAGGGTGAAGTGTGACCAAGTATTGG 4560
Qy 4561 CCTAGCAGAGCAGAGAAACCCACGGACTCGTTCAAGTAAAGCTGCTTTGATPACTGTGGAG 4620
Db 4561 CCTAGCAGAGCAGAGAAACCCACGGACTCGTTCAAGTAAAGCTGCTTTGATPACTGTGGAG 4620
Qy 4621 CTGGCCACATATTGTGTTTCGAACATTTTGCACTTTTCAAGAAATGGTTCAAGTGAAGAGAGA 4680
Db 4621 CTGGCCACATATTGTGTTTCGAACATTTTGCACTTTTCAAGAAATGGTTCAAGTGAAGAGAGA 4680
Qy 4681 GAAGTGAGACAATTCAGTTCCAGCTCAGCCGCTGATCATGTGTGTTCCAGAACACCCCTACA 4740
Db 4681 GAAGTGAGACAATTCAGTTCCAGCTCAGCCGCTGATCATGTGTGTTCCAGAACACCCCTACA 4740
Qy 4741 CCTTTTCTAGCTTTCTTACGTAGAGTCAAAACCTGTAAACCTCCCGATGCTGGTCCGATG 4800
Db 4741 CCTTTTCTAGCTTTCTTACGTAGAGTCAAAACCTGTAAACCTCCCGATGCTGGTCCGATG 4800
Qy 4801 GTTGTGCACTGAGTGGGAGTTGGCCGAGTGGTTCCTCATGCTCATAGATGCCATG 4860
Db 4801 GTTGTGCACTGAGTGGGAGTTGGCCGAGTGGTTCCTCATGCTCATAGATGCCATG 4860
Qy 4861 TTAGAAAGATTAAGCATGAAATACTGTAGATATTTATGCGCATGTAACTTTAATGAGA 4920
Db 4861 TTAGAAAGATTAAGCATGAAATACTGTAGATATTTATGCGCATGTAACTTTAATGAGA 4920
Qy 4921 CCCAGAGGAATATATGTTTCAACAGAGAACCAATACATCTTTATCCATGATGCACTG 4980
Db 4921 CCCAGAGGAATATATGTTTCAACAGAGAACCAATACATCTTTATCCATGATGCACTG 4980
Qy 4981 TTAGAGCAGTGACTTTGGGAATACCGAAGTCCGAGCTAGAACTGTATGCTCATATT 5040
Db 4981 TTAGAGCAGTGACTTTGGGAATACCGAAGTCCGAGCTAGAACTGTATGCTCATATT 5040
Qy 5041 CAGAAGCTGACACAAATAGAAACGGGAGAGAAATGTACAGGAATGGAGCTCGAATTTAAG 5100
Db 5041 CAGAAGCTGACACAAATAGAAACGGGAGAGAAATGTACAGGAATGGAGCTCGAATTTAAG 5100
Qy 5101 CGTCTAGCCAGCTCAAAAGCTCACACCTCAAGGTTTATCAGTGCGAATCTTTCCATGTAAT 5160
Db 5101 CGTCTAGCCAGCTCAAAAGCTCACACCTCAAGGTTTATCAGTGCGAATCTTTCCATGTAAT 5160
Qy 5161 AAATTCAAAATCGCTTTGTTAATATTTATGCCATATGAATCCACAGGGTATGCTGCGAG 5220
Db 5161 AAATTCAAAATCGCTTTGTTAATATTTATGCCATATGAATCCACAGGGTATGCTGCGAG 5220
Qy 5221 CCTATCCGTGGAGTAGAAGGATCTGATTAATACATCAATGCGAGTTTATTTGATGGATACAGA 5280
Db 5221 CCTATCCGTGGAGTAGAAGGATCTGATTAATACATCAATGCGAGTTTATTTGATGGATACAGA 5280
Qy 5281 CAACAGAAAGCTTACATCGCTACCCAGGGGCCCTTTGGCAGAGACCACTGAAGACTCTGCG 5340
Db 5281 CAACAGAAAGCTTACATCGCTACCCAGGGGCCCTTTGGCAGAGACCACTGAAGACTCTGCG 5340
Qy 5341 CGGATGCTCTGGGACACAATTCACCACTAGTTGATGCTCACAAGCTCGGTGAATG 5400
Db 5341 CGGATGCTCTGGGACACAATTCACCACTAGTTGATGCTCACCAGCTCGGTGAATG 5400
Qy 5401 GGCAGAGAGAAATGTCAACAATCTGCGCAGCAGAAACGGTCTGCAAGATACCAAGTACTTTT 5460
|||||

Db 5401 GGCAGAGAGAAATGTCAACAATCTGCGCCAGCAGAAACGGTCTGCAAGATACCAAGTACTTTT 5460
Qy 5461 GTTGTAGATCCCATGGCTGAGTACAAATGCGCAAGTATATCTTAAGGGAATTTCAAGGTC 5520
Db 5461 GTTGTAGATCCCATGGCTGAGTACAAATGCGCAAGTATATCTTAAGGGAATTTCAAGGTC 5520
Qy 5521 ACAGATCCAGGGAGCGCCAGTCCCGAAACAGTAGAGGAGTTCACAGTTCTACGTGCTGGCCA 5580
Db 5521 ACAGATCCAGGGAGCGCCAGTCCCGAAACAGTAGAGGAGTTCACAGTTCTACGTGCTGGCCA 5580
Qy 5581 GAGCAAGAGTGCCTCAAAAGTCCGGAAGAGGATTTATTGACTTTCATCGCGCAAGTCCATAAA 5640
Db 5581 GAGCAAGAGTGCCTCAAAAGTCCGGAAGAGGATTTATTGACTTTCATCGCGCAAGTCCATAAA 5640
Qy 5641 ACAAAGAAACAGTTTGGCCAAAGATGGACCCATTTTCACTTCAGTCCAGCGCGGGGCTTTGGA 5700
Db 5641 ACAAAGAAACAGTTTGGCCAAAGATGGACCCATTTTCACTTCAGTCCAGCGCGGGGCTTTGGA 5700
Qy 5701 AGAAGTGGAGTCTTCAATAACGCTAAGCATTTGTTTGGAAAGAAATGAGATATGAAGGAGTT 5760
Db 5701 AGAAGTGGAGTCTTCAATAACGCTAAGCATTTGTTTGGAAAGAAATGAGATATGAAGGAGTT 5760
Qy 5761 GTAGATATCTTCCAGACTGTCAAAATGTTAAGAACACAAACGAGCTATGGTACAGACA 5820
Db 5761 GTAGATATCTTCCAGACTGTCAAAATGTTAAGAACACAAACGAGCTATGGTACAGACA 5820
Qy 5821 GAGGATCAATATCAGTTTCTATCGTGGCGCACTAGAGTACCTGGGAGCTTTGACCCAC 5880
Db 5821 GAGGATCAATATCAGTTTCTATCGTGGCGCACTAGAGTACCTGGGAGCTTTGACCCAC 5880
Qy 5881 TATGCAACGTAAGAAACCCCTGACCCATTTCTGGATTTTCTACAGGCGCTTCAATATCCA 5940
Db 5881 TATGCAACGTAAGAAACCCCTGACCCATTTCTGGATTTTCTACAGGCGCTTCAATATCCA 5940
Qy 5941 TGGAGTCTTCTGAGCCATACAGGGCCTTTGAGAACTCTTTTAACTTCTAGCTAACA 6000
Db 5941 TGGAGTCTTCTGAGCCATACAGGGCCTTTGAGAACTCTTTTAACTTCTAGCTAACA 6000
Qy 6001 ACTACTTAGTGGGACTATTACACAAACAAATTAACAAATTTTCCAGCTGGAGCC 6060
Db 6001 ACTACTTAGTGGGACTATTACACAAACAAATTAACAAATTTTCCAGCTGGAGCC 6060
Qy 6061 AAGAAATCTTTGACATCGCCCTTCCCAACCATCTGCTCATATAAATTTAGGGGCCA 6120
Db 6061 AAGAAATCTTTGACATCGCCCTTCCCAACCATCTGCTCATATAAATTTAGGGGCCA 6120
Qy 6121 AGGGAGGGAATGTTTAAAAAGAAAGTCTTTGATTTAGTTTATTTAGTTTAAAGATAC 6180
Db 6121 AGGGAGGGAATGTTTAAAAAGAAAGTCTTTGATTTAGTTTATTTAGTTTAAAGATAC 6180
Qy 6181 TGCTGACCTGCTTCAATTTCTAACTGTAAACCTTTTAAACAAATGATCATTCG 6240
Db 6181 TGCTGACCTGCTTCAATTTCTAACTGTAAACCTTTTAAACAAATGATCATTCG 6240
Qy 6241 ATAAAGTGAATTTTAAAAAGTT 6263
Db 6241 ATAAAGTGAATTTTAAAAAGTT 6263

RESULT 2
US-09-023-655-879
; Sequence 879, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 5, 2005, 22:42:42 ; Search time 3361 Seconds
(without alignments)
11687.798 Million cell updates/sec

Title: US-10-772-636-63
Perfect score: 6263
Sequence: 1 gctaactcaaggagagcgc.....aagtgaatttttaaaagttt 6263

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 6313374 seqs, 3136092125 residues

Total number of hits satisfying chosen parameters: 12626748

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/FCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10J_PUBCOMB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6263	100.0	6263	18	US-10-641-643-1473
2	6263	100.0	6263	21	US-10-772-636-63
3	2473	39.5	6363	18	US-10-641-643-879
4	2359	37.7	7741	17	US-10-291-265-14
5	2357.4	37.6	7702	11	US-09-719-272-3
6	2357.4	37.6	7702	17	US-10-210-838-4
7	2357.4	37.6	7702	18	US-10-641-643-1336

8	2357.4	37.6	7702	21	US-10-945-647-3	Sequence 3, Appli
9	2357.4	37.6	7702	21	US-10-482-029-51	Sequence 51, Appli
10	2357.4	37.6	7702	21	US-10-945-772-3	Sequence 3, Appli
11	2357.4	37.6	7702	21	US-10-505-680-391	Sequence 391, App
12	2357.4	37.6	7718	14	US-10-175-523-192	Sequence 192, App
13	2357.4	37.6	7718	17	US-10-133-937-9	Sequence 9, Appli
14	2357.4	37.6	7718	17	US-10-159-563-9	Sequence 9, Appli
15	2357.4	37.6	7724	17	US-10-210-838-18	Sequence 18, Appli
16	2357.4	37.6	7945	20	US-10-357-930-27910	Sequence 27910, A
17	2348	37.5	7705	17	US-10-291-265-486	Sequence 486, App
18	2307.2	36.8	6734	21	US-10-764-420-2695	Sequence 2695, Ap
19	2303.2	36.8	8232	14	US-10-198-846-10976	Sequence 10976, A
20	1345.8	21.5	4078	17	US-10-305-720-1132	Sequence 1132, Ap
21	1290	20.6	6143	17	US-10-210-838-11	Sequence 11, Appli
22	1245.2	19.9	5412	17	US-10-258-666-11	Sequence 11, Appli
23	1196.6	19.1	4190	17	US-10-388-934-253	Sequence 253, App
24	1095.8	17.5	3467	11	US-09-719-272-1	Sequence 1, Appli
25	1095.8	17.5	3467	21	US-10-945-647-1	Sequence 1, Appli
26	1095.8	17.5	3467	21	US-10-945-772-1	Sequence 1, Appli
27	829.4	13.2	3064	9	US-09-925-300-10	Sequence 10, Appli
28	442	7.1	2872	21	US-10-777-144-4	Sequence 4, Appli
29	442	7.1	2872	21	US-10-777-145-4	Sequence 4, Appli
30	442	7.1	2872	21	US-10-777-186-4	Sequence 2, Appli
31	438.8	7.0	2409	21	US-10-777-144-2	Sequence 2, Appli
32	438.8	7.0	2409	21	US-10-777-145-2	Sequence 2, Appli
33	438.8	7.0	2430	17	US-10-210-556-227	Sequence 227, App
34	438.8	7.0	2430	17	US-10-210-556-226	Sequence 4701, Ap
35	438.8	7.0	2526	21	US-10-956-157-4701	Sequence 1311, Ap
36	438.8	7.0	2565	17	US-10-305-720-1311	Sequence 21, Appli
37	438.8	7.0	2565	17	US-10-210-556-21	Sequence 963, App
38	438.8	7.0	2826	17	US-10-264-049-963	Sequence 4, Appli
39	438.8	7.0	3148	17	US-10-210-556-4	Sequence 226, App
40	438.8	7.0	3279	17	US-10-210-556-226	Sequence 1254, App
41	438.8	7.0	3615	17	US-10-305-720-1254	Sequence 18, Appli
42	438.8	7.0	3615	17	US-10-210-556-18	Sequence 225, App
43	438.8	7.0	3643	17	US-10-210-556-225	Sequence 947, App
44	438.8	7.0	3643	21	US-10-956-157-947	Sequence 948, App
45	438.8	7.0	3643	21	US-10-956-157-948	

ALIGNMENTS

RESULT 1

US-10-641-643-1473
; Sequence 1473, Application US/10641643
; Publication No. US20040077003A1

GENERAL INFORMATION:

APPLICANT: Cocks, Benjamin G.

Susan G. Stuart

Jeffrey J. Seilhamer

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL

GENE EXPRESSION

NUMBER OF SEQUENCES: 1508

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/641.643

FILING DATE: 14-Aug-2003

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: <Unknown>

FILING DATE: <Unknown>

```
;
;
; ATTORNEY/AGENT INFORMATION:
;   NAME: Zeller, Karen J.
;   REGISTRATION NUMBER: 37,071
;   REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (650) 855-0555
;   TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1473:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 6263 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
; IMMEDIATE SOURCE:
;   LIBRARY: GENBANK
;   CLONE: 9755652
; SEQUENCE DESCRIPTION: SEQ ID NO: 1473 :
US-10-641-643-1473

Query Match      100.08; Score 6263; DB 18; Length 6263;
Best Local Similarity 100.08; Pred. No. 0;
Matches 6263; Conserved 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTAACTCAAGGAGACGCTCTGGTGAACACCCGCTGGGATCTAAAGAACAAAGCTCTGAAAG 60
DB 1 GCTAACTCAAGGAGACGCTCTGGTGAACACCCGCTGGGATCTAAAGAACAAAGCTCTGAAAG 60
QY 61 TGTTCAGCTGAAATTTTCAGATCGGACAGACTCGCTGCGGTCCGGAGGCGAGTATCCCA 120
DB 61 TGTTCAGCTGAAATTTTCAGATCGGACAGACTCGCTGCGGTCCGGAGGCGAGTATCCCA 120
QY 121 AGCTGCTCGGCACGCTCTGCCAGCTGCGAGATGGTGCAGCTAGCAGGCTCTGCTG 180
DB 121 AGCTGCTCGGCACGCTCTGCCAGCTGCGAGATGGTGCAGCTAGCAGGCTCTGCTG 180
QY 181 CTGCTCTCTCACTTTCTCCGACGAGATGCTGAGACACCTCCAAAGTTTACACGAACA 240
DB 181 CTGCTCTCTCACTTTCTCCGACGAGTGTGAGACACCTCCAAAGTTTACACGAACA 240
QY 241 CCGGTGATACAGACAGGGGTCTTGGCGGAGTGTGCTCTTTTCATCTGCAAGCTACGGGA 300
DB 241 CCGGTGATACAGACAGGGGTCTTGGCGGAGTGTGCTCTTTTCATCTGCAAGCTACGGGA 300
QY 301 GACCCAGACCTTAAATTTGCTGGGAACAAAGAGGAGAGTGCAGCAATCAGAGATT 360
DB 301 GACCCAGACCTTAAATTTGCTGGGAACAAAGAGGAGAGTGCAGCAATCAGAGATT 360
QY 361 GAGGTAAATAGAGTTTGACGATGGGTCTGGATCAGTTCTCAGAAATACACCCCTTACGGACT 420
DB 361 GAGGTAAATAGAGTTTGACGATGGGTCTGGATCAGTTCTCAGAAATACACCCCTTACGGACT 420
QY 421 CCGAGGATAGAGCCATTATGATGTGGCTCAAAATATGTGGGAGAAATAGTGTA 480
DB 421 CCGAGGATAGAGCCATTATGATGTGGCTCAAAATATGTGGGAGAAATAGTGTA 480
QY 481 TCCACAGACTCAGAGTTTTCGGGAAGATCAAAATTCACGGGCTCCCTACCATTTGAC 540
DB 481 TCCACAGACTCAGAGTTTTCGGGAAGATCAAAATTCACGGGCTCCCTACCATTTGAC 540
QY 541 ATGGGCCCAACAGTTGAAGGTGTTGAGCGTACTCGCACGGCCACCATCTTTGTGACGCC 600
DB 541 ATGGGCCCAACAGTTGAAGGTGTTGAGCGTACTCGCACGGCCACCATCTTTGTGACGCC 600
QY 601 AGTGTAAATCCGGATCCAGAAATCACTTGGTTTAAAGATTTCTTACCTGTGGACACAAGC 660
DB 601 AGTGTAAATCCGGATCCAGAAATCACTTGGTTTAAAGATTTCTTACCTGTGGACACAAGC 660
QY 661 AACACAAATGTCGTTAATAGCAGTTACGATCAGAAATCTATTGGTGTGACACCAATAGA 720
DB 661 AACACAAATGTCGTTAATAGCAGTTACGATCAGAAATCTATTGGTGTGACACCAATAGA 720
QY 721 GGAGCCCTTCAGATTGAGCAGAGTGAAGATCTGACCAAGGAAATATGAGTGTGTGCC 780
DB 721 GGAGCCCTTCAGATTGAGCAGAGTGAAGATCTGACCAAGGAAATATGAGTGTGTGCC 780
QY 781 ACCAACAGCGCGGCACTCGCTATTCCGCTCTCGCCTCCGCAATTTATATGTCAGAGAGCTCGGA 840
DB 781 ACCAACAGCGCGGCACTCGCTATTCCGCTCTCGCCTCCGCAATTTATATGTCAGAGAGCTCGGA 840
QY 841 GAAGTTGCGCGTGTCCACCAAGATTTCTATCCACCCCACTAATCATGAAATCATGCCA 900
DB 841 GAAGTTGCGCGTGTCCACCAAGATTTCTATCCACCCCACTAATCATGAAATCATGCCA 900
QY 901 GCGGAAGCGTTAATATCACCTGTGTGGCGGTGACCAATGCTTATGTAAGTGG 960
DB 901 GCGGAAGCGTTAATATCACCTGTGTGGCGGTGACCAATGCTTATGTAAGTGG 960
QY 961 ATGTTGGGGCAGAGATCTGACACCTGAAGATGATATGCCAATAGGAAGAAATGTGCTA 1020
DB 961 ATGTTGGGGCAGAGATCTGACACCTGAAGATGATATGCCAATAGGAAGAAATGTGCTA 1020
QY 1021 GAACTGAATGATGTAAGACAGTCTAGCAAAATTTACACCTGTGTGCTATGTCAACACTGGGT 1080
DB 1021 GAACTGAATGATGTAAGACAGTCTAGCAAAATTTACACCTGTGTGCTATGTCAACACTGGGT 1080
QY 1081 GTCAATTGAAGCAATAGCACAGATCACTGTCAAAAGCCTTACCACCACTCCAGGAATCTCT 1140
DB 1081 GTCAATTGAAGCAATAGCACAGATCACTGTCAAAAGCCTTACCACCACTCCAGGAATCTCT 1140
QY 1141 GTAGTGACCGAGACAGCTCAAGCATCACACTGACGTGGGACTCTCGGGAACCCCTGAG 1200
DB 1141 GTAGTGACCGAGACAGCTCAAGCATCACACTGACGTGGGACTCTCGGGAACCCCTGAG 1200
QY 1201 CCTGTTTCTTATTACATAAATTCAGCATATAAACTCTGAGGAACCTTTACAAAGAA 1260
DB 1201 CCTGTTTCTTATTACATAAATTCAGCATATAAACTCTGAGGAACCTTTACAAAGAA 1260
QY 1261 ATTGATGGGTGGCGACCAACACGCTACAGTGTGCTGGAATAAGTCTCTACTCGGATAT 1320
DB 1261 ATTGATGGGTGGCGACCAACACGCTACAGTGTGCTGGAATAAGTCTCTACTCGGATAT 1320
QY 1321 GAATTCAGGGTTGTTGCTGCTCAATATGTTGGGGGGGGCTCCACAGGAACCTGTGCTA 1380
DB 1321 GAATTCAGGGTTGTTGCTGCTCAATATGTTGGGGGGGGCTCCACAGGAACCTGTGCTA 1380
QY 1381 ACACAAACCTCAGAGCAAGCAACATCCAGTGTCCCGAGGGATGTCCAGGCAAGATGTTG 1440
DB 1381 ACACAAACCTCAGAGCAAGCAACATCCAGTGTCCCGAGGGATGTCCAGGCAAGATGTTG 1440
QY 1441 AGTTCGACCAACATTTTGGTACAGTGGAAAGAACTCTGAAAGCAAAATGGACAGATCAA 1500
DB 1441 AGTTCGACCAACATTTTGGTACAGTGGAAAGAACTCTGAAAGCAAAATGGACAGATCAA 1500
QY 1501 GGATATAGAGTTTATATACAATGGATCCCACTCAACATGTCAACACTGGATGAAACAC 1560
DB 1501 GGATATAGAGTTTATATACAATGGATCCCACTCAACATGTCAACACTGGATGAAACAC 1560
QY 1561 AATGTAGCTGACAGCCAAATCACTACTATTGGCAACTTAGTGCCCCAGAAAACATATTTCT 1620
DB 1561 AATGTAGCTGACAGCCAAATCACTACTATTGGCAACTTAGTGCCCCAGAAAACATATTTCT 1620
QY 1621 GTCAAAAGTCTGGCTTTTACTCAATTTGGAGATGGTCCCTTTCAAGTGCATACAGTGC 1680
DB 1621 GTCAAAAGTCTGGCTTTTACTCAATTTGGAGATGGTCCCTTTTCAAGTGCATACAGTGC 1680
QY 1681 ATCACTCAGACAGAGTACCAAGGCGAGCCACTAACTTCAAGCAGAACTCTGAGTCTGAA 1740
DB 1681 ATCACTCAGACAGAGTACCAAGGCGAGCCACTAACTTCAAGCAGAACTCTGAGTCTGAA 1740
QY 1741 ACAAGTATTTTGTCTCTCTTGGACACCTCCAGTTTCAGATACCAATGCAACTG 1800
DB 1741 ACAAGTATTTTGTCTCTCTTGGACACCTCCAGTTTCAGATACCAATGCAACTG 1800
QY 1801 GTCTAAGAGATGGGAGCATGGAGAGGAGCAAGAAATACCATTTGACCCAGGACATCA 1860
DB 1801 GTCTAAGAGATGGGAGCATGGAGAGGAGCAAGAAATACCATTTGACCCAGGACATCA 1860
```

Qy	1861	TATAGGCTGCAAGGACTGAAACCAACAGCTTATATCTATTTCCGCTCGCTGCAAGCTCC	1920
Db	1861	TATAGGCTGCAAGGACTGAAACCAACAGCTTATATCTATTTCCGCTCGCTGCAAGCTCC	1920
Qy	1921	CCTCAAGGCTGGGTCTTCTACTCGAAGATATCAGCTAGAACCATCGAGTCAAGCGG	1980
Db	1921	CCTCAAGGCTGGGTCTTCTACTCGAAGATATCAGCTAGAACCATCGAGTCAAGCGG	1980
Qy	1981	TCAGCTCTCTCAAGACATTTAGTTGCAACCAAGTTCCACCTAGTATTTTGGTAAAGT	2040
Db	1981	TCAGCTCTCTCAAGACATTTAGTTGCAACCAAGTTCCACCTAGTATTTTGGTAAAGT	2040
Qy	2041	TGGCAACCTCCACAGTGGGAAACAGAAATGSCATTTATCTGAATATCTGCATCAAGTAC	2100
Db	2041	TGGCAACCTCCACAGTGGGAAACAGAAATGSCATTTATCTGAATATCTGCATCAAGTAC	2100
Qy	2101	ACTGCAGTGGATGGGAAAGATGACAAAGCTCAGAGATTTTGGGAATTTCCCTCGGACACT	2160
Db	2101	ACTGCAGTGGATGGGAAAGATGACAAAGCTCAGAGATTTTGGGAATTTCCCTCGGACACT	2160
Qy	2161	ACCAATACCTTTTGGAAACAGCTGGAAATATGGAATACCGGATCACTGTGACAGCC	2220
Db	2161	ACCAATACCTTTTGGAAACAGCTGGAAATATGGAATACCGGATCACTGTGACAGCC	2220
Qy	2221	CATACAGATGTCGGCCCTGGCCCTGAGAGCTTGTCCGTGTTGATTCGAAACCAATGAAGAT	2280
Db	2221	CATACAGATGTCGGCCCTGGCCCTGAGAGCTTGTCCGTGTTGATTCGAAACCAATGAAGAT	2280
Qy	2281	GTTCCCTAGTGGTCTCTCGAAAGCTCGAGGTAGAGGCTGCAACTCAACATCTGTTAAA	2340
Db	2281	GTTCCCTAGTGGTCTCTCGAAAGCTCGAGGTAGAGGCTGCAACTCAACATCTGTTAAA	2340
Qy	2341	GTCTCATGGCGCTCACCCGTCGCCAATAAACAGCATGGCCAGATAGAGGATATCAGGTG	2400
Db	2341	GTCTCATGGCGCTCACCCGTCGCCAATAAACAGCATGGCCAGATAGAGGATATCAGGTG	2400
Qy	2401	CATTATGTAGAGATGGGAAATTTGAGCGCCAAAGGCCAGGCCCATGCTGAAAGATGTCATG	2460
Db	2401	CATTATGTAGAGATGGGAAATTTGAGCGCCAAAGGCCAGGCCCATGCTGAAAGATGTCATG	2460
Qy	2461	CTGGCTGATGCACAGTGGGAAATTTGATGATCTACTGAAACATGACATCATTTCTGGG	2520
Db	2461	CTGGCTGATGCACAGTGGGAAATTTGATGATCTACTGAAACATGACATCATTTCTGGG	2520
Qy	2521	CTCCAGCCTGAAACTTCTACTCTCCCTCACCGCTCACAGCCTACACAAACCAAGGAGTGGT	2580
Db	2521	CTCCAGCCTGAAACTTCTACTCTCCCTCACCGCTCACAGCCTACACAAACCAAGGAGTGGT	2580
Qy	2581	GCTCGCAGAACCCAAACTGGTGTCCACACTGGGGCAGTTCACAGGAAACCTCGGCTT	2640
Db	2581	GCTCGCAGAACCCAAACTGGTGTCCACACTGGGGCAGTTCACAGGAAACCTCGGCTT	2640
Qy	2641	GTGATTAACCACTCAGATGAATCTCTTATTCAGTGGCACTCCGCTGGACACA	2700
Db	2641	GTGATTAACCACTCAGATGAATCTCTTATTCAGTGGCACTCCGCTGGACACA	2700
Qy	2701	TTTGGACCTCTTCAGGGCTACCGTCTAAAATTTGGCCGCAAGGATATGGAGCCACTTACT	2760
Db	2701	TTTGGACCTCTTCAGGGCTACCGTCTAAAATTTGGCCGCAAGGATATGGAGCCACTTACT	2760
Qy	2761	ACTCTTGAGTCTCTGAAACAGAGATCACTTTTACAGCTACAGACATCCACAGGGAGCA	2820
Db	2761	ACTCTTGAGTCTCTGAAACAGAGATCACTTTTACAGCTACAGACATCCACAGGGAGCA	2820
Qy	2821	TCATACGCTCTCAGGCTCTCAGCAGAAACAAAGTGGCTTTTGGGGAGGATGTTGAAG	2880
Db	2821	TCATACGCTCTCAGGCTCTCAGCAGAAACAAAGTGGCTTTTGGGGAGGATGTTGAAG	2880
Qy	2881	GAGATTTCCATTTCCAGAGAGATACCAACTGGATTTCCCTCAAAACCTTCACTCAGAGGC	2940
Db	2881	GAGATTTCCATTTCCAGAGAGATACCAACTGGATTTCCCTCAAAACCTTCACTCAGAGGC	2940

Qy	2941	ACCATTCTCAACCTCCGCTCAGTTATCTTTGGCAACCACTGCTCTGGCAGAGAGAAATGGC	3000
Db	2941	ACCATTCTCAACCTCCGCTCAGTTATCTTTGGCAACCACTGCTCTGGCAGAGAGAAATGGC	3000
Qy	3001	ATTATCAACCAAGTATACCTCTTTTATAGGATATCAACATCCCTCTCCCGATGGAG	3060
Db	3001	ATTATCAACCAAGTATACCTCTTTTATAGGATATCAACATCCCTCTCCCGATGGAG	3060
Qy	3061	CAGCTTATTTGTCAGCTGCACCACTATGACACTCACTGGCTTTAAACCCAGATACCAACA	3120
Db	3061	CAGCTTATTTGTCAGCTGCACCACTATGACACTCACTGGCTTTAAACCCAGATACCAACA	3120
Qy	3121	TACGATGTAAAGTACGCTGCTCATACGAGCAAGGGCCCGGCCATATAGTCCGAGTGC	3180
Db	3121	TACGATGTAAAGTACGCTGCTCATACGAGCAAGGGCCCGGCCATATAGTCCGAGTGC	3180
Qy	3181	CAGTTTCAGGACACTGCTGCTGATCAAGTGTTCGAAAAATTTTCATGTCAAAGCAGTA	3240
Db	3181	CAGTTTCAGGACACTGCTGCTGATCAAGTGTTCGAAAAATTTTCATGTCAAAGCAGTA	3240
Qy	3241	ATGAAGACTTCCGCTGCTGCTCTTGGGAGATTCAGAGAAATTAATCTCCGCCATGCCT	3300
Db	3241	ATGAAGACTTCCGCTGCTGCTCTTGGGAGATTCAGAGAAATTAATCTCCGCCATGCCT	3300
Qy	3301	TTCAAAATTTCTTATGATGATGGGAAATGGTAGAAGTGGATGGCCGAGCCACACAG	3360
Db	3301	TTCAAAATTTCTTATGATGATGGGAAATGGTAGAAGTGGATGGCCGAGCCACACAG	3360
Qy	3361	AAGTTAATTTGCAACCTGMAAGCCTGAGAAATCATATTCATTTGCTGACAAAATCGTGA	3420
Db	3361	AAGTTAATTTGCAACCTGMAAGCCTGAGAAATCATATTCATTTGCTGACAAAATCGTGA	3420
Qy	3421	AACAGTCTGGTGGCTGTCAGCACAGGGTCAACGGCAAGAGCTGCACACAGATGTATTACGT	3480
Db	3421	AACAGTCTGGTGGCTGTCAGCACAGGGTCAACGGCAAGAGCTGCACACAGATGTATTACGT	3480
Qy	3481	ACCAAGCTGCTTCATTTGGGAGAACCACTTGGATGGCATGATCTCTGTGCAACTGCCT	3540
Db	3481	ACCAAGCTGCTTCATTTGGGAGAACCACTTGGATGGCATGATCTCTGTGCAACTGCCT	3540
Qy	3541	GAAGTACCTGCAAAATGAGAAATATAAAGTTTACTACATAATAATTTGTCCTTTGAAGAAA	3600
Db	3541	GAAGTACCTGCAAAATGAGAAATATAAAGTTTACTACATAATAATTTGTCCTTTGAAGAAA	3600
Qy	3601	TCTCGCGGAAATTTTCAAGCCATGGGAGAGTCCAGATGAAATGGAATTAAGATGAGCTG	3660
Db	3601	TCTCGCGGAAATTTTCAAGCCATGGGAGAGTCCAGATGAAATGGAATTAAGATGAGCTG	3660
Qy	3661	CTTAAGGAGATATCTAGAAAGCCGAGAAAGCATCCGTTATGGGAGAGAAATTAAG	3720
Db	3661	CTTAAGGAGATATCTAGAAAGCCGAGAAAGCATCCGTTATGGGAGAGAAATTAAG	3720
Qy	3721	CCATATATTCGCGCTCACTTTGATGCTCTTCCCACTGAGTTTCACTGGGGGATGACAAAG	3780
Db	3721	CCATATATTCGCGCTCACTTTGATGCTCTTCCCACTGAGTTTCACTGGGGGATGACAAAG	3780
Qy	3781	CATTATGCTGGATTTCAAAACCAAGCACTCCAAAGTGGTCAAGAAATATGTCTTCTTGTG	3840
Db	3781	CATTATGCTGGATTTCAAAACCAAGCACTCCAAAGTGGTCAAGAAATATGTCTTCTTGTG	3840
Qy	3841	TTAGCAGTAAATGGAACATGCAAGAGTCTAAGATGATGCAACAGCCCTTACTCCGACCCC	3900
Db	3841	TTAGCAGTAAATGGAACATGCAAGAGTCTAAGATGATGCAACAGCCCTTACTCCGACCCC	3900
Qy	3901	GTGTGTCAATGGATCTGGATCCGACCAATCAGGATGAAGAAAGAGCTTCATCTGG	3960
Db	3901	GTGTGTCAATGGATCTGGATCCGACCAATCAGGATGAAGAAAGAGCTTCATCTGG	3960
Qy	3961	GTTGTAGGCTCTGCTTGTGAGTGGTCTTTTATCATCTGCATTTGTCAATTCCTTCTT	4020
Db	3961	GTTGTAGGCTCTGCTTGTGAGTGGTCTTTTATCATCTGCATTTGTCAATTCCTTCTT	4020
Qy	4021	TATAAAGGAAGGGGAGAGTGGGACTCTAGAAAAAGCAGCATACCCGAACAATAAGAG	4080

Db	4021		TATAAAAGGAAGAGGCGAGAGTCCGACTCTAGAAAAAGCAGCATACCGAAACAATAAGGAG	4080
Qy	4081	ATCCCTTCACACCAACCCAAACAGACCTGTAGAACTGAGGCGCTTAACTTTCAACACCCG	4140	
Db	4081	ATCCCTTCACACCAACCCAAACAGACCTGTAGAACTGAGGCGCTTAACTTTCAACACCCG	4140	
Qy	4141	GGTATGGCTAGCCATCTCTCAAATACCCATCTTTGGAACTTCGAGACCAATTTGAAAGATTG	4200	
Db	4141	GGTATGGCTAGCCATCTCTCAAATACCCATCTTTGGAACTTCGAGACCAATTTGAAAGATTG	4200	
Qy	4201	AAAGCAATATGACAACTTTGAAAGTTTCCAGGAATATGAGTCAATTTGACCTGTGCCACGAG	4260	
Db	4201	AAAGCAATATGACAACTTTGAAAGTTTCCAGGAATATGAGTCAATTTGACCTGTGCCACGAG	4260	
Qy	4261	TTTCACTTTGGGAACATTCAAACTTTGGAGTAAACAAACCAAGATATAGATACGGCAATGTA	4320	
Db	4261	TTTCACTTTGGGAACATTCAAACTTTGGAGTAAACAAACCAAGATATAGATACGGCAATGTA	4320	
Qy	4321	ATCCATATATGATCATCTCCGGGTTCTCCTATCAGCTATAGAAGGGATCCCGAGAAATGAC	4380	
Db	4321	ATCCATATATGATCATCTCCGGGTTCTCCTATCAGCTATAGAAGGGATCCCGAGAAATGAC	4380	
Qy	4381	TATGTGAATGCCAACTACATAGATGGGTATAGGAAGCAAAATGCTTATTTGCAACACAG	4440	
Db	4381	TATGTGAATGCCAACTACATAGATGGGTATAGGAAGCAAAATGCTTATTTGCAACACAG	4440	
Qy	4441	GGATCTCTCCCCGAAAAATTTGGGGACCTTTTGGGAATGATATGGGAACCAACCGAGTGCC	4500	
Db	4441	GGATCTCTCCCCGAAAAATTTGGGGACCTTTTGGGAATGATATGGGAACCAACCGAGTGCC	4500	
Qy	4501	ACAGTTGTATGATGACAAAACCTAGAAAGATCAAGGGTGAAGTGTGACCAAGTATTGG	4560	
Db	4501	ACAGTTGTATGATGACAAAACCTAGAAAGATCAAGGGTGAAGTGTGACCAAGTATTGG	4560	
Qy	4561	CCTAGCAGAGGCCACAGAAAACCCAGGACTCGTTCAAAGTAAACCGTCTGATCTGTGGAG	4620	
Db	4561	CCTAGCAGAGGCCACAGAAAACCCAGGACTCGTTCAAAGTAAACCGTCTGATCTGTGGAG	4620	
Qy	4621	CTGGCCACATATTTGTTTCGNAACATTTGCACCTTTTACAAGATGGTTCAGGTGAGAGAGA	4680	
Db	4621	CTGGCCACATATTTGTTTCGNAACATTTGCACCTTTTACAAGATGGTTCAGGTGAGAGAGA	4680	
Qy	4681	GAAGTGAGACAAATTTCCAGTTTCAACCGCTGGGCTCATCATGGTGTTCAGAAACACCTTACA	4740	
Db	4681	GAAGTGAGACAAATTTCCAGTTTCAACCGCTGGGCTCATCATGGTGTTCAGAAACACCTTACA	4740	
Qy	4741	CCTTTTCTAGCTTTCTTAAGTAGAGTCAAAAACCTGTAAACCTCCCGATGTGTCGATG	4800	
Db	4741	CCTTTTCTAGCTTTCTTAAGTAGAGTCAAAAACCTGTAAACCTCCCGATGTGTCGATG	4800	
Qy	4801	GTTGTGCACTCAGTGGGGAGTTGGCCGAGCTGGTCTTCATCGTACATAGTGCCATG	4860	
Db	4801	GTTGTGCACTCAGTGGGGAGTTGGCCGAGCTGGTCTTCATCGTACATAGTGCCATG	4860	
Qy	4861	TTAGAAAGAAATAAGCATGAAAAAATCTAGATATTTATGGCCATGTAACTTTTAATGAGA	4920	
Db	4861	TTAGAAAGAAATAAGCATGAAAAAATCTAGATATTTATGGCCATGTAACTTTTAATGAGA	4920	
Qy	4921	GCCCAGAGGAACCTATATGGTTCAAAACAGAAACCAATACATCTTTATCCATGTGCACTG	4980	
Db	4921	GCCCAGAGGAACCTATATGGTTCAAAACAGAAACCAATACATCTTTATCCATGTGCACTG	4980	
Qy	4981	TTAGAAAGCAGTGACTTTGGGAAATACCGAAGTGGCAGCTAGAAACTTGTATGCCCTACATT	5040	
Db	4981	TTAGAAAGCAGTGACTTTGGGAAATACCGAAGTGGCAGCTAGAAACTTGTATGCCCTACATT	5040	
Qy	5041	CAGAAAGCTGACACAAAATAGAAACCGGAGAGAAATGTCAAGAAATGGAGCTCGAATTTAAG	5100	
Db	5041	CAGAAAGCTGACACAAAATAGAAACCGGAGAGAAATGTCAAGAAATGGAGCTCGAATTTAAG	5100	
Qy	5101	CGTCTAGCCAGCTCAAAAGCTCACACCTCAAGGTTTATCAGTGGCAATCTTCCATGTAAAT	5160	

5101	Db	CGTCTAGCCAGCTCAAAGAGCTCACACTCAAGGTTTATCAGTGCCTCAATCTTCCATGTAAAT	5161
5161	Qy	AAATTCAAATAATCGCCTTGTTAAATATATATGCCATATGAATCCACAAGGGGTATGCCTGCAG	5220
5161	Db	AAATTCAAATAATCGCCTTGTTAAATATATGCCATATGAATCCACAAGGGGTATGCCTGCAG	5220
5221	Qy	CCTATCCGTGGAGTAGAAGGATCTGATTAACATCAATGCCCAGTTTATTTGATGATACAGACA	5280
5221	Db	CCTATCCGTGGAGTAGAAGGATCTGATTAACATCAATGCCCAGTTTATTTGATGATACAGACA	5280
5281	Qy	CACAGAAAGCCTACATCGCTACCCAGGGGCCCTTCGCAGAGACCACTGAAGACTTCTCTGG	5340
5281	Db	CACAGAAAGCCTACATCGCTACCCAGGGGCCCTTCGCAGAGACCACTGAAGACTTCTCTGG	5340
5341	Qy	CGGATGCTCTGGGAAACAACAATTCACACATAGTGTGTGATGCTCACCAAGCTGCGTGAATG	5400
5341	Db	CGGATGCTCTGGGAAACAACAATTCACACATAGTGTGTGATGCTCACCAAGCTGCGTGAATG	5400
5401	Qy	GGCAGAGAGAAATGTCAACCAATACTGGCCAGCAGAAACGGTCTGCAAGAATACAGTACTTTT	5460
5401	Db	GGCAGAGAGAAATGTCAACCAATACTGGCCAGCAGAAACGGTCTGCAAGAATACAGTACTTTT	5460
5461	Qy	GTTGTAGATCCCATGCTGAGTACACATGCCACAGTATATCCTTAAGGGAATTCAAAGTCT	5520
5461	Db	GTTGTAGATCCCATGCTGAGTACACATGCCACAGTATATCCTTAAGGGAATTCAAAGTCT	5520
5521	Qy	ACAGATGCCAGGACGGCCAGTCCCGAAACAGTAAGGCAGTTCACAGTTTCACTGACTGGCCA	5580
5521	Db	ACAGATGCCAGGACGGCCAGTCCCGAAACAGTAAGGCAGTTCACAGTTTCACTGACTGGCCA	5580
5581	Qy	GAGCAAGGAGTGCCAAAGTCCGAGAAAGGATTTATTGAATTCATTCGGCCCAAGTCCATAAA	5640
5581	Db	GAGCAAGGAGTGCCAAAGTCCGAGAAAGGATTTATTGAATTCATTCGGCCCAAGTCCATAAA	5640
5641	Qy	ACAAAAGACAGTTTGGCCAGAGATGACCCATTTTCAGTCCATTCGACGGCGGGCGTTTGG	5700
5641	Db	ACAAAAGACAGTTTGGCCAGAGATGACCCATTTTCAGTCCATTCGACGGCGGGCGTTTGG	5700
5701	Qy	AGAACTGGAGTCTTCATACGCTTAAGCATTTGTTTGGAAAGATGAGATATGAAGGAGTT	5760
5701	Db	AGAACTGGAGTCTTCATACGCTTAAGCATTTGTTTGGAAAGATGAGATATGAAGGAGTT	5760
5761	Qy	GTAGATATCTTCAGACTGTCAAATGTTTAAGAACACAACGACAGCTATGGTACAGACA	5820
5761	Db	GTAGATATCTTCAGACTGTCAAATGTTTAAGAACACAACGACAGCTATGGTACAGACA	5820
5821	Qy	GAGGATCAATATCAGTTTTCCTATCGTCCGCACTAGAGTACCTTGGCAGCTTTGACCA	5880
5821	Db	GAGGATCAATATCAGTTTTCCTATCGTCCGCACTAGAGTACCTTGGCAGCTTTGACCA	5880
5881	Qy	TATGCAACGTAGAAAACCCCTGACCCATTCGTGATTTTTTACTACAGGCCCTTCAATATCCA	5940
5881	Db	TATGCAACGTAGAAAACCCCTGACCCATTCGTGATTTTTTACTACAGGCCCTTCAATATCCA	5940
5941	Qy	TGGAGTCTCTTCTGAGCCATACAGGGCATTGAGAACTCCTTCTTAACCTTCTAGCTAAACA	6000
5941	Db	TGGAGTCTCTCTTCTGAGCCATACAGGGCATTGAGAACTCCTTCTTAACCTTCTAGCTAAACA	6000
6001	Qy	ACTACTTAGTGGGACTATTACACAAAAACAATTTAAAAACAATTTTCCAGGTGACCC	6060
6001	Db	ACTACTTAGTGGGACTATTACACAAAAACAATTTAAAAACAATTTTCCAGGTGACCC	6060
6061	Qy	AAGAAATCTTTTGTACATCGCCCTTCCACCATCTGCTCATATAACATTTTATAGGGGCCA	6120
6061	Db	AAGAAATCTTTTGTACATCGCCCTTCCACCATCTGCTCATATAACATTTTATAGGGGCCA	6120
6121	Qy	AGGGGAGGGAATGTTTAAAAAAGAAAGTCTTGTATTTAGTTTTTGTGTAAGATAC	6180
6121	Db	AGGGGAGGGAATGTTTAAAAAAGAAAGTCTTGTATTTAGTTTTTGTGTAAGATAC	6180
6181	Qy	TGCTGACTGTGCTTCAATTTCTAACTGTGTAAACCTTTTTTTTTTAAACAAAATGTATCATTCG	6240
6181	Db	TGCTGACTGTGCTTCAATTTCTAACTGTGTAAACCTTTTTTTTTTAAACAAAATGTATCATTCG	6240


```
Qy 6241 ATAAAGTGAATTTTAAAAAGTT 6263
Db 6241 ATAAAGTGAATTTTAAAAAGTT 6263

RESULT 2
US-10-772-636-63
; Sequence 63, Application US/10772636
; Publication No. US20050042687A1
; GENERAL INFORMATION:
; APPLICANT: Kelly, Louise M.
; APPLICANT: Carroll, Joseph M.
; APPLICANT: Farlow, Deborah
; APPLICANT: Healy, Aileen
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: HEMATOLOGICAL DISORDERS USING 9118, 990, 17662, 81982, 630,
; TITLE OF INVENTION: 21472, 17692, 19290, 21620, 21689, 28899, 53659, 64549,
; TITLE OF INVENTION: 9465, 23544, 7366, 27417, 57259, 21844, 943, 2061, 5891,
; TITLE OF INVENTION: 9137, 13908, 14310, 17600, 25584, 27824, 28469, 38947,
; TITLE OF INVENTION: 53003, 965, 56639, 9661, 16052, 1521, 6662, 13913, 12405 OR
; TITLE OF INVENTION: 5014
; FILE REFERENCE: MPI03-01SP1RNMNM
; CURRENT APPLICATION NUMBER: US/10/772,636
; CURRENT FILING DATE: 2004-02-05
; PRIOR FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US 60/445,241
; PRIOR FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US 60/448,389
; PRIOR FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 60/456,320
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 60/460,279
; PRIOR FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: US 60/465,924
; PRIOR FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: US 60/470,052
; PRIOR FILING DATE: 2003-05-13
; PRIOR APPLICATION NUMBER: US 60/498,106
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: US 60/500,179
; PRIOR FILING DATE: 2003-09-04
; PRIOR APPLICATION NUMBER: US 60/502,909
; PRIOR FILING DATE: 2003-09-15
; PRIOR APPLICATION NUMBER: US 60/510,351
; PRIOR FILING DATE: 2003-10-10
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 6263
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (154)...(5892)
US-10-772-636-63

Query Match 100.0%; Score 6263; DB 21; Length 6263;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTAACTCAAGGAGAGCTCTGGTGAACACCCGTGGGATCTAAAGAACAAAGCTCTGAAAG 60
Db 1 GCTAACTCAAGGAGAGCTCTGGTGAACACCCGTGGGATCTAAAGAACAAAGCTCTGAAAG 60

Qy 61 TGTTCAGCTCAAAATTCAGATCGGACAGACTCCGCTCGGGCTCCGGAGGAGTGAATCCA 120
Db 61 TGTTCAGCTCAAAATTCAGATCGGACAGACTCCGCTCGGGCTCCGGAGGAGTGAATCCA 120

Qy 121 AGCTGCTCGGCACGCTCTGCTCCAAAGTGTGACAGTGTGACAGGATGTGACAGTGTGCTG 180
Db 121 AGCTGCTCGGCACGCTCTGCTCCAAAGTGTGACAGTGTGACAGGATGTGACAGTGTGCTG 180
```

```
Qy 181 CTGCTCCTCACTTCTCTCTCCGCA CGGATGCTGAGACACCTCCAAGTTTACACGAACA 240
Db 181 CTGCTCCTCACTTCTCTCTCCGCA CGGATGCTGAGACACCTCCAAGTTTACACGAACA 240

Qy 241 CCGTTGATCAGACAGGGGTCTCTGGCGGAGTTCCTCTCTTTTCATCTGCAAGCTACGGGA 300
Db 241 CCGTTGATCAGACAGGGGTCTCTGGCGGAGTTCCTCTCTTTTCATCTGCAAGCTACGGGA 300

Qy 301 GACCCAAAGACTAAATTTGTCTGGAACAAAAAGAAAGTACAGCAATCAGAGATTT 360
Db 301 GACCCAAAGACTAAATTTGTCTGGAACAAAAAGAAAGTACAGCAATCAGAGATTT 360

Qy 361 GAGGTAATAGAGTTTGAGATGGGTCTGGATCAGTCTCAGAAATACAACTTACGAGCT 420
Db 361 GAGGTAATAGAGTTTGAGATGGGTCTGGATCAGTCTCAGAAATACAACTTACGAGCT 420

Qy 421 CCGAGGATCAGGGCCATTTTATGAATGTGCGCTCAAAATATGTGGGAGAAATAGTGTA 480
Db 421 CCGAGGATCAGGGCCATTTTATGAATGTGCGCTCAAAATATGTGGGAGAAATAGTGTA 480

Qy 481 TCACCAAGACTCAGATTTTTCGGGAAAGATCAAAATCCCAAGGGCTTCCCTACCAATGAC 540
Db 481 TCACCAAGACTCAGATTTTTCGGGAAAGATCAAAATCCCAAGGGCTTCCCTACCAATGAC 540

Qy 541 ATGGGCCCCACAGTTGAAGTGGTTGACGCTACTCGCAGCGCCACCATTGCTTGTGACGC 600
Db 541 ATGGGCCCCACAGTTGAAGTGGTTGACGCTACTCGCAGCGCCACCATTGCTTGTGACGC 600

Qy 601 AGTGGTAATCCGGATCCAGAAATCAGTTGGTTTAAAGATTTCTTACCTGTGGACACAGC 660
Db 601 AGTGGTAATCCGGATCCAGAAATCAGTTGGTTTAAAGATTTCTTACCTGTGGACACAGC 660

Qy 661 AACAACAATGGTGGTATTAAGCAGTTACGATCAGAAATCTATTTGGTGGTACACCAATAGA 720
Db 661 AACAACAATGGTGGTATTAAGCAGTTACGATCAGAAATCTATTTGGTGGTACACCAATAGA 720

Qy 721 GGAGCCCTCAGATTGAGCAGAGTGAAGAGTGTGACCAAGGAAATATGAGTGTGTGCC 780
Db 721 GGAGCCCTCAGATTGAGCAGAGTGAAGAGTGTGACCAAGGAAATATGAGTGTGTGCC 780

Qy 781 ACCAAGACGGGGGCACTCCCTATTCCGCTCCCTCCCAATTTATATGTCAGAGACTCGGA 840
Db 781 ACCAAGACGGGGGCACTCCCTATTCCGCTCCCTCCCAATTTATATGTCAGAGACTCGGA 840

Qy 841 GAAAGTTCGCGGTGTCCCAACCAAGATTTCTATCCCAACCACTAATCATGAATCATGCCA 900
Db 841 GAAAGTTCGCGGTGTCCCAACCAAGATTTCTATCCCAACCACTAATCATGAATCATGCCA 900

Qy 901 GCGGGAAGCGTTAATATCACTGTGTGGCGGTGCGGTGACCAATGCCCTTATGTAAGTGG 960
Db 901 GCGGGAAGCGTTAATATCACTGTGTGGCGGTGCGGTGACCAATGCCCTTATGTAAGTGG 960

Qy 961 ATGTTGGGGGACAGAGATCTGACACCTGAAGATGATATGCAATAGGAAGAAATGTGCTA 1020
Db 961 ATGTTGGGGGACAGAGATCTGACACCTGAAGATGATATGCAATAGGAAGAAATGTGCTA 1020

Qy 1021 GAACTGAATGATGTAAAGCAGTCAAGCAATTAACACCTGTGTGTGTATGTCAACTGGGT 1080
Db 1021 GAACTGAATGATGTAAAGCAGTCAAGCAATTAACACCTGTGTGTGTATGTCAACTGGGT 1080

Qy 1081 GTCAATGAAGCAATAGCAGATCACTGTCAAGACCTTACCACCACTCCAGGAACTCCT 1140
Db 1081 GTCAATGAAGCAATAGCAGATCACTGTCAAGACCTTACCACCACTCCAGGAACTCCT 1140

Qy 1141 GTAGTGACCGAGACAGCTCAAGCATCAGCTGCGGGGACTCTCGGAAACCCCTGAG 1200
Db 1141 GTAGTGACCGAGACAGCTCAAGCATCAGCTGCGGGGACTCTCGGAAACCCCTGAG 1200

Qy 1201 CTTGTTTCTTATTACATAAATTCAGCATAAACCTTAAAGCTCTGAGGAACCTTACAAAGAA 1260
Db 1201 CTTGTTTCTTATTACATAAATTCAGCATAAACCTTAAAGCTCTGAGGAACCTTACAAAGAA 1260

Qy 1261 ATTGATGGGGTGGGACCAACACGCTACAGTGTGCTGGAGCTAAGTCCCTACTCGGATAT 1320
```

1261 ATTGATGGGTGGGACACACGCTACAGTGTCCGTGCACTAAGTCCCTACTCGATTAT 1320
 1321 GAATTCAGGGTTGTTGCTGTCAATAAATTTGGGCGGGGCTCCAGCGAACTGTGCTA 1380
 1321 GAATTCAGGGTTGTTGCTGTCAATAAATTTGGGCGGGGCTCCAGCGAACTGTGCTA 1380
 1381 ACACAAACCTCAGAGCAAGCAATCCAGTGTCCCGGAGGATGTCCAGGCAAGATGTTG 1440
 1381 ACACAAACCTCAGAGCAAGCAATCCAGTGTCCCGGAGGATGTCCAGGCAAGATGTTG 1440
 1441 AGTTCGACCACTTTTGTGTACGTGGAAGCACTCAAGAGCCAAATGACAGATCCAA 1500
 1441 AGTTCGACCACTTTTGTGTACGTGGAAGCACTCAAGAGCCAAATGACAGATCCAA 1500
 1501 GGATATAGAGTTTATATACAATGGATCCCACTCAACATGTCAACAACTGGATGAACAC 1560
 1501 GGATATAGAGTTTATATACAATGGATCCCACTCAACATGTCAACAACTGGATGAACAC 1560
 1561 AATGTAGCTGACGCCAAATCACTACTATTGGGCAACTTAGTGCCCCAGAAACATATTCT 1620
 1561 AATGTAGCTGACGCCAAATCACTACTATTGGGCAACTTAGTGCCCCAGAAACATATTCT 1620
 1621 GTCAAGTCCCTGGCTTTTACCTCAATTTGGAGATGCTCCCTTTCAAGTGCATACAAGTC 1680
 1621 GTCAAGTCCCTGGCTTTTACCTCAATTTGGAGATGCTCCCTTTCAAGTGCATACAAGTC 1680
 1681 ATCACTCAGACAGGAGTACAGGGCAGCCACTAACTTTCAAAGCAGAACTGTGTA 1740
 1681 ATCACTCAGACAGGAGTACAGGGCAGCCACTAACTTTCAAAGCAGAACTGTGTA 1740
 1741 ACAAGTATTTTGTCTCTTTGACACCTTCCAGTTCAGATACCATTTGCAACTATGAAGTC 1800
 1741 ACAAGTATTTTGTCTCTTTGACACCTTCCAGTTCAGATACCATTTGCAACTATGAAGTC 1800
 1801 GTCTACAAGATGGGAGCATGGAGAGGAGCAACGAATTTACATTTGAGCCAGGACATCA 1860
 1801 GTCTACAAGATGGGAGCATGGAGAGGAGCAACGAATTTACATTTGAGCCAGGACATCA 1860
 1861 TATAGGCTGCAAGGACTGAAACCAAAACAGCTTATATTTCCGTCTGGCTGCAAGTC 1920
 1861 TATAGGCTGCAAGGACTGAAACCAAAACAGCTTATATTTCCGTCTGGCTGCAAGTC 1920
 1921 CCTCAAGGCTGGGTGCTTCTACTGCAAGATATCAGCTAGAACCATTCAGTCAAGGCG 1980
 1921 CCTCAAGGCTGGGTGCTTCTACTGCAAGATATCAGCTAGAACCATTCAGTCAAGGCG 1980
 1981 TCAGTCTCTCTCAAGACATTTAGTTGCAACGAGCCCAAGTTCACCTAGTATTTTGTGAAT 2040
 1981 TCAGTCTCTCTCAAGACATTTAGTTGCAACGAGCCCAAGTTCACCTAGTATTTTGTGAAT 2040
 2041 TGGCAACCTCCACAGTGGGAAAACAGAAATGGCAATTTACTGTAATCTCCATCAAGTAC 2100
 2041 TGGCAACCTCCACAGTGGGAAAACAGAAATGGCAATTTACTGTAATCTCCATCAAGTAC 2100
 2101 ACTCAGTGGATGGGAAAGATGAACGCTCAAGATTTTGGGAAATTCCTTGGGACACT 2160
 2101 ACTCAGTGGATGGGAAAGATGAACGCTCAAGATTTTGGGAAATTCCTTGGGACACT 2160
 2161 ACCAAATACCTTTTGGAAACAGTGGGAAAATGGCAATACCGGATCAGTGCAGGCC 2220
 2161 ACCAAATACCTTTTGGAAACAGTGGGAAAATGGCAATACCGGATCAGTGCAGGCC 2220
 2221 CATACAGATGTCCGCGCTGGGCTGAGAGCTTGTCCGTGTTGATTCGAAACCAATGAAGAT 2280
 2221 CATACAGATGTCCGCGCTGGGCTGAGAGCTTGTCCGTGTTGATTCGAAACCAATGAAGAT 2280
 2281 GTTCTAGTGGTCTCTCGAAAGTCCAGGTAGAGGCTGTCAACTCAACATCTGTTAAA 2340
 2281 GTTCTAGTGGTCTCTCGAAAGTCCAGGTAGAGGCTGTCAACTCAACATCTGTTAAA 2340
 2341 GTCTCATGGCTCACCCGTGCCCAATAAAGCATGGCCAGATGAAGGATATCAGGTG 2400

2341 GTCTCATGGCTCACCCGTGCCAATAAAGCATGGCCAGATAAGAGGATATCAGGTG 2400
 2401 CATTATGTGAGGATGAAATGGTGAAGCCCAAGGGCAGCCCATGCTGAAAGATGTCATG 2460
 2401 CATTATGTGAGGATGAAATGGTGAAGCCCAAGGGCAGCCCATGCTGAAAGATGTCATG 2460
 2461 CTGGCTGATGCAAGTGGGAAATTTGATGATACCTACTACTGTAACATGACATGATCAIT 2520
 2461 CTGGCTGATGCAAGTGGGAAATTTGATGATACCTACTACTGTAACATGACATGATCAIT 2520
 2521 CTCACGCTGAAATCTTCTACTCCCTCAGCGTCAAGCTTACACAAAGAGAGATGTT 2580
 2521 CTCACGCTGAAATCTTCTACTCCCTCAGCGTCAAGCTTACACAAAGAGAGATGTT 2580
 2581 GCTGCGAGCAAGCCCAAACTGGTGTCCACACTGGGGCAGTTCAGGGAACCTCGGCTT 2640
 2581 GCTGCGAGCAAGCCCAAACTGGTGTCCACACTGGGGCAGTTCAGGGAACCTCGGCTT 2640
 2641 GTGATTTAACCACTCAGATGAATATCTGCTCTTATTTCAGTGGCACCTCCGGTGGACACA 2700
 2641 GTGATTTAACCACTCAGATGAATATCTGCTCTTATTTCAGTGGCACCTCCGGTGGACACA 2700
 2701 TTTGGACCTCTTCCAGGCTACCGTCTAAATTTGGCGCAAGATATGGAGCCACTTACT 2760
 2701 TTTGGACCTCTTCCAGGCTACCGTCTAAATTTGGCGCAAGATATGGAGCCACTTACT 2760
 2761 ACTCTTGGATTTCTGTAAGAAAGATCACTTTTACAGCTTACAGATCCACAGGGAGCA 2820
 2761 ACTCTTGGATTTCTGTAAGAAAGATCACTTTTACAGCTTACAGATCCACAGGGAGCA 2820
 2821 TCATACGCTCTTCCAGGCTCTCAGCCAGAAACAAAGTGGGCTTTGGGAGGAGATGGTGAAG 2880
 2821 TCATACGCTCTTCCAGGCTCTCAGCCAGAAACAAAGTGGGCTTTGGGAGGAGATGGTGAAG 2880
 2881 GAGATTTCCATTTCCAGAGAAAGTACCAACTGGATTCCTCAAAACCTTCACTCAGAAAGGC 2940
 2881 GAGATTTCCATTTCCAGAGAAAGTACCAACTGGATTCCTCAAAACCTTCACTCAGAAAGGC 2940
 2941 ACCACTTCAACCTCCGTCAGTTTATCTTGGCAACCACTGTCTCGCAGAGAGAAATGGC 3000
 2941 ACCACTTCAACCTCCGTCAGTTTATCTTGGCAACCACTGTCTCGCAGAGAGAAATGGC 3000
 3001 ATTATCAACCAAGTATACCTCTTTTATAGGGAATCAACATCCCTCTCCCGATGGAG 3060
 3001 ATTATCAACCAAGTATACCTCTTTTATAGGGAATCAACATCCCTCTCCCGATGGAG 3060
 3061 CAGCTTATTTCTCCAGCTGACACCTATGACACTCACTGGCTTAAACCAAGATACACACA 3120
 3061 CAGCTTATTTCTCCAGCTGACACCTATGACACTCACTGGCTTAAACCAAGATACACACA 3120
 3121 TAGATGTAAGATACGCTGTCTATACGAGCAAGGGCCCGGCGCATATAGTCCCGATGTC 3180
 3121 TAGATGTAAGATACGCTGTCTATACGAGCAAGGGCCCGGCGCATATAGTCCCGATGTC 3180
 3181 CAGTTCAAGACACTGCTGCTGGATCAAGTGTGCAAAAAATTTTCAATGTCAAAAGCAGTA 3240
 3181 CAGTTCAAGACACTGCTGCTGGATCAAGTGTGCAAAAAATTTTCAATGTCAAAAGCAGTA 3240
 3241 ATGAAGACTTCCGTTGCTGTCTTGGGAGATTCAGAGAAATTAATACTCCGCTATGCT 3300
 3241 ATGAAGACTTCCGTTGCTGTCTTGGGAGATTCAGAGAAATTAATACTCCGCTATGCT 3300
 3301 TTCAAAATTTCTTATGATGATGGGAAATGGTAGAAGTGGATGGCCGAGCCACACAG 3360
 3301 TTCAAAATTTCTTATGATGATGGGAAATGGTAGAAGTGGATGGCCGAGCCACACAG 3360
 3361 AAGTTAATTTGTCAACCTGAGAAATCAATTTTCTGCTGACAAATCGTGA 3420
 3361 AAGTTAATTTGTCAACCTGAGAAATCAATTTTCTGCTGACAAATCGTGA 3420
 3421 AACAGTCTGGTGGCTGCAGCACAGGCTCACGCAAAAGACTGCACAGATGTATTACGT 3480
 3421 AACAGTCTGGTGGCTGCAGCACAGGCTCACGCAAAAGACTGCACAGATGTATTACGT 3480

QY 3481 ACCAAGCCTGCTTCACTTGGGAAGACCAACTTGGATGGCATGATTAATCTGTGCAACTGCT 3540
DB |||||
QY 3481 ACCAAGCCTGCTTCACTTGGGAAGACCAACTTGGATGGCATGATTAATCTGTGCAACTGCT 3540
DB |||||
QY 3541 GAAGTACCTGCAAAATGAGAATATAAAGGTTACTACATAAATTAATTTGCTTTGAAGAA 3600
DB |||||
QY 3541 GAAGTACCTGCAAAATGAGAATATAAAGGTTACTACATAAATTAATTTGCTTTGAAGAA 3600
DB |||||
QY 3601 TCTCGGGGAAATTTATCAAGCCATGGGAGATCCAGATGAAATGAAATAGATGAGCTG 3660
DB |||||
QY 3601 TCTCGGGGAAATTTATCAAGCCATGGGAGATCCAGATGAAATGAAATAGATGAGCTG 3660
DB |||||
QY 3661 CTTAAGGAGATATCTAGGAAGCGCAAGAGCATCCGTTATGGAGAGAGTTGAAATTAAG 3720
DB |||||
QY 3661 CTTAAGGAGATATCTAGGAAGCGCAAGAGCATCCGTTATGGAGAGAGTTGAAATTAAG 3720
DB |||||
QY 3721 CCATATATTTGCCGCTCACCTTTGATGCTCTTCCACCTGAGTTTCCCTGGGGATGACAAG 3780
DB |||||
QY 3721 CCATATATTTGCCGCTCACCTTTGATGCTCTTCCACCTGAGTTTCCCTGGGGATGACAAG 3780
DB |||||
QY 3781 CATATGTTGGATTTACAAACAGCAACTCCAAAGTGGTCAAGAAATATGTCTTTTGTG 3840
DB |||||
QY 3781 CATATGTTGGATTTACAAACAGCAACTCCAAAGTGGTCAAGAAATATGTCTTTTGTG 3840
DB |||||
QY 3841 TTAGCAGTAATGGAACATGCGAGTCTAAGATGTATGCAACAGCCCTTATCTCGACCCC 3900
DB |||||
QY 3841 TTAGCAGTAATGGAACATGCGAGTCTAAGATGTATGCAACAGCCCTTATCTCGACCCC 3900
DB |||||
QY 3901 GTGGTGTCAATGGATCTCGATCCGAGCCCATCAAGATGAGAGAGAGGCTTGATCTGG 3960
DB |||||
QY 3901 GTGGTGTCAATGGATCTCGATCCGAGCCCATCAAGATGAGAGAGAGGCTTGATCTGG 3960
DB |||||
QY 3961 GTTGTAGTCTGCTCTTGGAGTGTCTTTATCATCTGCAATTTGATGCTATTTCTT 4020
DB |||||
QY 3961 GTTGTAGTCTGCTCTTGGAGTGTCTTTATCATCTGCAATTTGATGCTATTTCTT 4020
DB |||||
QY 4021 TATATAAGGAGAGGGCAGAGTCCGACTCTAGAAAAAGCAGCATACCCGAAACAATAAGGAG 4080
DB |||||
QY 4021 TATATAAGGAGAGGGCAGAGTCCGACTCTAGAAAAAGCAGCATACCCGAAACAATAAGGAG 4080
DB |||||
QY 4081 ATCCCTTCCACACCCACAGACCCCTGTAGACTGAGGCGCTTAACTTTCAACACCG 4140
DB |||||
QY 4081 ATCCCTTCCACACCCACAGACCCCTGTAGACTGAGGCGCTTAACTTTCAACACCG 4140
DB |||||
QY 4141 GGTATGGCTAGCCATCTCCAAATACCCATCTTGGAACTTGGAACTTGGAACTTGGAACTTGG 4200
DB |||||
QY 4141 GGTATGGCTAGCCATCTCCAAATACCCATCTTGGAACTTGGAACTTGGAACTTGGAACTTGG 4200
DB |||||
QY 4201 AAAGCAAAATGCAACTTGAAGTTTCCAGGAATATGATCAATTTGACCCCTGGCCAGCAG 4260
DB |||||
QY 4201 AAAGCAAAATGCAACTTGAAGTTTCCAGGAATATGATCAATTTGACCCCTGGCCAGCAG 4260
DB |||||
QY 4261 TTCACTTGGGAACATTTCAAACTTGGAACTTGGAACTTGGAACTTGGAACTTGGAACTTGG 4320
DB |||||
QY 4261 TTCACTTGGGAACATTTCAAACTTGGAACTTGGAACTTGGAACTTGGAACTTGGAACTTGG 4320
DB |||||
QY 4321 ATCGCATATGATCAATTTCCGGGTTCTCTATCAGCTATAGAGGATCCAGGAGTGAC 4380
DB |||||
QY 4321 ATCGCATATGATCAATTTCCGGGTTCTCTATCAGCTATAGAGGATCCAGGAGTGAC 4380
DB |||||
QY 4381 TATGTGAATGCAACTACATAGATGGGTATAGGAAGCAAAATGCTTATTTGCAACACAG 4440
DB |||||
QY 4381 TATGTGAATGCAACTACATAGATGGGTATAGGAAGCAAAATGCTTATTTGCAACACAG 4440
DB |||||
QY 4441 GATCTCTCCCGGAAACATTTGGGACCTTTTGGAGAAATGATATGGGAAACAACGGAGTGCC 4500
DB |||||
QY 4441 GATCTCTCCCGGAAACATTTGGGACCTTTTGGAGAAATGATATGGGAAACAACGGAGTGCC 4500
DB |||||
QY 4501 ACAGTTGTCTGATGACAAAACCTAGAAAGAGATCAAGGGTGAAGTGTGACCATTTGG 4560
DB |||||
QY 4501 ACAGTTGTCTGATGACAAAACCTAGAAAGAGATCAAGGGTGAAGTGTGACCATTTGG 4560
DB |||||

QY 4561 CCTAGCAGAGCACAAGAAACCCACGGAATCGTTCAAGTAAACGCTGCTGATCTGTGAG 4620
DB |||||
QY 4561 CCTAGCAGAGCACAAGAAACCCACGGAATCGTTCAAGTAAACGCTGCTGATCTGTGAG 4620
DB |||||
QY 4621 CTGGCCACATATTTGTTTCAAGAACTTTGCACTTTTACAGAAATGTTCAAGTGAGAAGAGA 4680
DB |||||
QY 4621 CTGGCCACATATTTGTTTCAAGAACTTTGCACTTTTACAGAAATGTTCAAGTGAGAAGAGA 4680
DB |||||
QY 4681 GAAGTGAGACAAATTTCCAGTTTCAAGCTGCTGCTGATCATTTGTTTCCAGAAACCCCTACA 4740
DB |||||
QY 4681 GAAGTGAGACAAATTTCCAGTTTCAAGCTGCTGCTGATCATTTGTTTCCAGAAACCCCTACA 4740
DB |||||
QY 4741 CCTTTTCTAGCTTTCTTACGTAGAGTCAAAACCTGTTAAACCTCCCGATGCTGCTCCGATG 4800
DB |||||
QY 4741 CCTTTTCTAGCTTTCTTACGTAGAGTCAAAACCTGTTAAACCTCCCGATGCTGCTCCGATG 4800
DB |||||
QY 4801 GTTGTGCACTCGAGTGGGGAGTTGGCGGAGTGGTGGTCTCATCTCATAGATGCCATG 4860
DB |||||
QY 4801 GTTGTGCACTCGAGTGGGGAGTTGGCGGAGTGGTGGTCTCATCTCATAGATGCCATG 4860
DB |||||
QY 4861 TTAGAAAGATTAAGAGCATGAAAAAATCTGTAGATATTTATGGCCATGTAACTTTTAAATGAGA 4920
DB |||||
QY 4861 TTAGAAAGATTAAGAGCATGAAAAAATCTGTAGATATTTATGGCCATGTAACTTTTAAATGAGA 4920
DB |||||
QY 4921 GCCCAGAGAACTATATGTTTCAAAACAGAGAACCAATACATCTTTATCCATGATGACATG 4980
DB |||||
QY 4921 GCCCAGAGAACTATATGTTTCAAAACAGAGAACCAATACATCTTTATCCATGATGACATG 4980
DB |||||
QY 4981 TTAGAAGCAGTACCTTTGTTGAAATACCGAAGTGGCAGTACAGAACTTGTATGCTTACATTT 5040
DB |||||
QY 4981 TTAGAAGCAGTACCTTTGTTGAAATACCGAAGTGGCAGTACAGAACTTGTATGCTTACATTT 5040
DB |||||
QY 5041 CAGAAAGTGAACAATAAGTAGAAACCGGAGAGAAATGTACAGGAATGGAGCTCGAAATTTAAG 5100
DB |||||
QY 5041 CAGAAAGTGAACAATAAGTAGAAACCGGAGAGAAATGTACAGGAATGGAGCTCGAAATTTAAG 5100
DB |||||
QY 5101 CGTCTAGCAGCTCAAAAGCTCACACTCAAGGTTTATCAGTGCCAACTTTCCATGTAAT 5160
DB |||||
QY 5101 CGTCTAGCAGCTCAAAAGCTCACACTCAAGGTTTATCAGTGCCAACTTTCCATGTAAT 5160
DB |||||
QY 5161 AAATTTCAAAATCCCTTTGTTTAAATATGATGCAATGATGATGATGATGATGATGATGATGATG 5220
DB |||||
QY 5161 AAATTTCAAAATCCCTTTGTTTAAATATGATGCAATGATGATGATGATGATGATGATGATGATG 5220
DB |||||
QY 5221 CCTATCGTGGAGTAGAAGATCTGATTTACATCAATGCGAGTTTATTTATGATGATGATGATGATG 5280
DB |||||
QY 5221 CCTATCGTGGAGTAGAAGATCTGATTTACATCAATGCGAGTTTATTTATGATGATGATGATGATG 5280
DB |||||
QY 5281 CAACAGAAAGCCTTACATGCTTACCCAGGGCCCTTGGCAGAGACACATGGAAGATCTTCTGG 5340
DB |||||
QY 5281 CAACAGAAAGCCTTACATGCTTACCCAGGGCCCTTGGCAGAGACACATGGAAGATCTTCTGG 5340
DB |||||
QY 5341 CGGATGCTCTGGGAACAAATTTCCACCCTAGTTGATGCTCACCAGCTCGCTGCTGCTGCTGCTGCTG 5400
DB |||||
QY 5341 CGGATGCTCTGGGAACAAATTTCCACCCTAGTTGATGCTCACCAGCTCGCTGCTGCTGCTGCTGCTG 5400
DB |||||
QY 5401 GGCAAGAGAAATGCTCACCCTTCTGGCCAGAGAGAGGCTCTGCAAGATACCAAGTACTTTT 5460
DB |||||
QY 5401 GGCAAGAGAAATGCTCACCCTTCTGGCCAGAGAGAGGCTCTGCAAGATACCAAGTACTTTT 5460
DB |||||
QY 5461 GTTGTAGATCCCATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5520
DB |||||
QY 5461 GTTGTAGATCCCATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5520
DB |||||
QY 5521 ACAGATCCAGGAGCGCCAGTCCCGAAGAGTGGAGTTTCCAGTTTCCAGTCTGCTGCTGCTGCTGCTG 5580
DB |||||
QY 5521 ACAGATCCAGGAGCGCCAGTCCCGAAGAGTGGAGTTTCCAGTTTCCAGTCTGCTGCTGCTGCTGCTG 5580
DB |||||
QY 5581 GAGCAAGAGTGGCCAAAGTCCGGAAGAGGATTTTATTTGACTTTTATCGGCAAGTCCATATAA 5640
DB |||||
QY 5581 GAGCAAGAGTGGCCAAAGTCCGGAAGAGGATTTTATTTGACTTTTATCGGCAAGTCCATATAA 5640
DB |||||
QY 5641 ACAAAGAAACAGTTTGGCCAGAGTGGACCCATTTTCACTGTCATGTCAGCGGGGCTTGGG 5700

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2005, 19:22:54 ; Search time 217 Seconds
(without alignments)
3407.769 Million cell updates/sec

Title: US-10-772-636-64

Perfect score: 10042

Sequence: 1 MVHVARLLLLLLFFLRLTDA.....YQFSYRAALEYLGSFPHYAT 1912

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_16Dec04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10042	100.0	1912	8	ADR40183 Human pro
2	7369.5	73.4	1949	7	ADE57117 Human pro
3	7369.5	73.4	1949	7	ADE57117 Human pro
4	7369.5	73.4	1949	7	ADD47015 Human pro
5	7369.5	73.4	1949	7	ADD47015 Human pro
6	7359.5	73.3	1948	7	ADD18742 Human dis
7	7245	72.1	1911	2	AAR71726 Human PTP
8	7245	72.1	1911	2	AAR7225 Human pro
9	7245	72.1	1911	2	AAW94027 Human pro
10	7245	72.1	1911	4	AAU01459 Human pro
11	7110	70.8	1904	5	ABBS7100 Mouse isc
12	7037	70.1	1907	4	AAU14143 Human nov
13	7032	70.0	1907	3	AAU14143 Human pro
14	7032	70.0	1907	3	AAU14143 Human pro
15	7032	70.0	1907	3	AAU14143 Human pro
16	7032	70.0	1907	3	AAU14143 Human pro
17	7032	70.0	1907	3	AAU14143 Human pro
18	7032	70.0	1907	3	AAU14143 Human pro
19	6977.5	69.5	1863	7	ADD46989 Rat Prote
20	6823.5	66.0	1786	6	AAE37971 Human kin
21	6308.5	62.8	1254	8	ADN02662 Liver dis
22	6251.5	62.3	1291	2	AAU75201 Tyrosine
23	7245.5	57.2	1501	2	AAU72858 Rat recep
24	5559.5	55.4	1495	5	ABBS7380 Rat mucoc
25	5559.5	55.4	1496	7	ADE57115 Rat Prote

26	5559.5	55.4	1496	7	ADE57119	Ades7119 Rat Prote
27	5559.5	55.4	1496	7	ADD47013	Add47013 Rat Prote
28	5559.5	55.4	1496	7	ADD47017	Add47017 Rat Prote
29	5333	53.1	1477	8	ADJ33671	Adj33671 Mouse leu
30	4865	48.4	2037	4	ABB71928	Abb71928 Drosophil
31	4664.5	46.4	1266	8	ADO66041	Adq66041 Novel hum
32	3655	36.4	960	8	ADR08858	Adr08858 Human pro
33	3027.5	30.1	647	4	AAU23746	Aam23746 Human EST
34	3027.5	30.1	647	4	AAU14379	Aau14379 Human nov
35	3027.5	30.1	647	8	ADH80697	Adh80697 Human pol
36	2965.5	29.5	1194	6	ABP97679	Abp97679 Amino aci
37	2926	29.1	607	3	AAU81783	Aay81783 Human pro
38	2926	29.1	607	3	AAU56098	Aay56098 LAR tyros
39	2238	22.3	1585	8	ADN22916	Adn22916 Bacterial
40	2238	22.3	1585	8	ADN22917	Adn22917 Bacterial
41	2166	21.6	442	3	AAU56372	Aab56372 Human pro
42	2024	20.2	608	8	ADR09270	Adr09270 Human pro
43	1957.5	19.5	400	2	AAU75203	Aar75203 Tyrosine
44	1814	18.1	344	7	ADJ69864	Adj69864 Human hea
45	1587	15.8	793	2	AAR20743	Aar20743 Murine re

ALIGNMENTS

RESULT 1

ADR40183
ID ADR40183 standard; protein; 1912 AA.

XX ADR40183;

AC AC

DT 18-NOV-2004 (first entry)

XX Human protein tyrosine phosphatase delta (R-PTP-delta) (965) protein.

XX haematological; cytostatic; erythroid; anaemia; erythrocytosis;

KW bone marrow; leukaemia; platelet; thrombocytopenia; thrombosis; B-cell;

KW T-cells; neutropenia; gene therapy; human;

KW protein tyrosine phosphatase delta; R-PTP-delta; enzyme.

XX Homo sapiens.

OS Homo sapiens.

XX WO2004072242-A2.

XX 26-AUG-2004.

XX 05-FEB-2004; 2004WO-US003417.

XX 05-FEB-2003; 2003US-0445241P.

XX 18-FEB-2003; 2003US-0448389P.

XX 20-MAR-2003; 2003US-0456320P.

XX 03-APR-2003; 2003US-0460279P.

XX 28-APR-2003; 2003US-0465924P.

XX 13-MAY-2003; 2003US-0470052P.

XX 26-AUG-2003; 2003US-0498106P.

XX 04-SEP-2003; 2003US-0500179P.

XX 15-SEP-2003; 2003US-0502909P.

XX 10-OCT-2003; 2003US-0510351P.

XX 17-OCT-2003; 2003US-0512380P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Kelly LM, Carroll JM, Farlow D, Healy A;

XX WPI; 2004-625850/60.

XX N-PSDB; ADR40182.

XX Identifying a compound capable of treating a hematological disorder

PT comprises combining a compound to be tested with a polypeptide related

PT with the disorder under conditions suitable for binding of the test

PT compound to the polypeptide.

XX Claim 1; SEQ ID NO 64; 321pp; English.

XX	The invention relates to a novel method for identifying a compound capable of treating a haematological disorder which comprises combining a compound to be tested with a specific polypeptide under conditions suitable for binding of the test compound to the polypeptide. The method of the invention has haematological and cytostatic applications and may be useful for identifying compounds for treating a haematological disorder associated with erythroid cells e.g. anaemia and erythrocytosis, bone marrow e.g. leukaemia, platelets e.g. thrombocytopenia and thrombosis or B-cells and T-cells e.g. neutropenia. The compounds identified may be utilised during gene therapy procedures. The current sequence is that of a human haematological disorder-related protein of the invention.				
XX	Sequence 1912 AA;				
QY	Query Match	100.0%;	Score 10042;	DB 8;	Length 1912;
DB	Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches 0;	Indels 0;
DB	Matches 1912;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MVHVARLLLLLLTFFLRDAETPRTRTPVDQTGSGGVASHFICQATGDPKPIVWKK	60		
DB	1	MVHVARLLLLLLTFFLRDAETPRTRTPVDQTGSGGVASHFICQATGDPKPIVWKK	60		
QY	61	GKGVSNQRFVIEFDDGSGSVLRIOPLRPRDEAIYECVASNNGEISVSTRITVLREDQ	120		
DB	61	GKGVSNQRFVIEFDDGSGSVLRIOPLRPRDEAIYECVASNNGEISVSTRITVLREDQ	120		
QY	121	IPRGFTIDMGPKLVKVERTRTATMLCAASGNPDPEITWFKDFLPVDTSNNGRIKQLRS	180		
DB	121	IPRGFTIDMGPKLVKVERTRTATMLCAASGNPDPEITWFKDFLPVDTSNNGRIKQLRS	180		
QY	181	ESIGGTPIRGALQIEQSESDQKCYCVATNSAGTRYSAPANLYVRELREVRVPPRFSI	240		
DB	181	ESIGGTPIRGALQIEQSESDQKCYCVATNSAGTRYSAPANLYVRELREVRVPPRFSI	240		
QY	241	PPTNHIMPVGSVNIITCVAVGSPMPYVKKWMLGAEITLTPDDMPIGRNVLLENDVROSANY	300		
DB	241	PPTNHIMPVGSVNIITCVAVGSPMPYVKKWMLGAEITLTPDDMPIGRNVLLENDVROSANY	300		
QY	301	TCVAMSTLGVIEAIIQITVKALPKPGTPVVTSTATSLTLDWDSGNPEPVSYYIIQHKP	360		
DB	301	TCVAMSTLGVIEAIIQITVKALPKPGTPVVTSTATSLTLDWDSGNPEPVSYYIIQHKP	360		
QY	361	KNSEELYKEIDGATTRYVAGLSPSYDFEFRVAVVNNIGRGPSPPEVLTQTSEQAPSSA	420		
DB	361	KNSEELYKEIDGATTRYVAGLSPSYDFEFRVAVVNNIGRGPSPPEVLTQTSEQAPSSA	420		
QY	421	PRDQARMLSSITLIVQWKEPEPNQIOGYRVYYTMDPTQHVNNMKHNVADSQITTTIG	480		
DB	421	PRDQARMLSSITLIVQWKEPEPNQIOGYRVYYTMDPTQHVNNMKHNVADSQITTTIG	480		
QY	481	NLPVQKTSYVKVLAFTSISGDPLSSDIQVITQTGVPQPLNFKAPESETSILLSWTPPR	540		
DB	481	NLPVQKTSYVKVLAFTSISGDPLSSDIQVITQTGVPQPLNFKAPESETSILLSWTPPR	540		
QY	541	SDTIANYELVYKDGEGEQRITIEPGTSYRLQGLKPNLSIYFRLAARSPQGLGASTAEI	600		
DB	541	SDTIANYELVYKDGEGEQRITIEPGTSYRLQGLKPNLSIYFRLAARSPQGLGASTAEI	600		
QY	601	SARTMQSKSAPPQDISCTSPSSTSLVSWQPPPEVKQNGIITEYSIKYTAVDGEDDKPH	660		
DB	601	SARTMQSKSAPPQDISCTSPSSTSLVSWQPPPEVKQNGIITEYSIKYTAVDGEDDKPH	660		
QY	661	EILGIPSDTTKYLLEQLEKWEYRIITVTAHTDVGPSPESLSVLIRTNEDVSPGPPKVEV	720		
DB	661	EILGIPSDTTKYLLEQLEKWEYRIITVTAHTDVGPSPESLSVLIRTNEDVSPGPPKVEV	720		
QY	721	EAVNSTSVKWSRSPVFNKHQGIQIRGVYVYVMEWENGEKGPMLKDVMLADAQWFEFDDT	780		
DB	721	EAVNSTSVKWSRSPVFNKHQGIQIRGVYVYVMEWENGEKGPMLKDVMLADAQWFEFDDT	780		
QY	781	TEHDMIISGLQETSYSYSLTVTAATTKGDGARSKPKLVSTTGAVPGKPRLVINHNTQNTAL	840		

781	TEHDMIISGLQETSYSYSLTVTAATTKGDGARSKPKLVSTTGAVPGKPRLVINHNTQNTAL	840	
841	IQWHPVDVTFGLQYRLKFGKQMEPLTTLBFSKEDHFTATDTHKGASVYFRLSARK	900	
841	IQWHPVDVTFGLQYRLKFGKQMEPLTTLBFSKEDHFTATDTHKGASVYFRLSARK	900	
901	VGFGEWKEIISPREVPTGPPQNLHSGTSTVSQLSWQPPVLAERNGLIITKYLTYRD	960	
901	VGFGEWKEIISPREVPTGPPQNLHSGTSTVSQLSWQPPVLAERNGLIITKYLTYRD	960	
961	INIPLLPMEQLIVPADTTMTLTGLKPDPTDYDVKVAHSTSKGPGYSPSQVFTLPVDQVF	1020	
961	INIPLLPMEQLIVPADTTMTLTGLKPDPTDYDVKVAHSTSKGPGYSPSQVFTLPVDQVF	1020	
1021	AKNFHVAVMKTSLVLSWEIPENYNSAMPFKILYDDGKMAVEVDGRATOKLIVNLKPEKS	1080	
1021	AKNFHVAVMKTSLVLSWEIPENYNSAMPFKILYDDGKMAVEVDGRATOKLIVNLKPEKS	1080	
1081	YSFVLTNRGSAGGLOHRTAKTAPDLRTKPAFTGKTNLDGMITVQLPEVPANENIKGY	1140	
1081	YSFVLTNRGSAGGLOHRTAKTAPDLRTKPAFTGKTNLDGMITVQLPEVPANENIKGY	1140	
1141	YIIIVPLKKSRCFKIPWESPDMEDELDELLEISKRKRSIRYGREVELKPYIAAHFDVLP	1200	
1141	YIIIVPLKKSRCFKIPWESPDMEDELDELLEISKRKRSIRYGREVELKPYIAAHFDVLP	1200	
1201	TEFTLGDDKHGGFTNKKQSGQEVYFVFLAWEHAESKMYATSPYSDPVWMDLDPQPI	1260	
1201	TEFTLGDDKHGGFTNKKQSGQEVYFVFLAWEHAESKMYATSPYSDPVWMDLDPQPI	1260	
1261	TDDEEGLTWWGVPVLAUVFIIICIVTALLYKRAESDSRKSSIINNKEIPSHHPTDPVE	1320	
1261	TDDEEGLTWWGVPVLAUVFIIICIVTALLYKRAESDSRKSSIINNKEIPSHHPTDPVE	1320	
1321	LRLNFQTPGMASHPPPIPIELADHIERLKANDNLKFSQYESIDPGQQFTWEHNSLEVN	1380	
1321	LRLNFQTPGMASHPPPIPIELADHIERLKANDNLKFSQYESIDPGQQFTWEHNSLEVN	1380	
1381	KPKNYANVIAVDHRSVLLSAIEGIPGSDYNVANYIDGVRKQNAVYATOGSIPETFGDFW	1440	
1381	KPKNYANVIAVDHRSVLLSAIEGIPGSDYNVANYIDGVRKQNAVYATOGSIPETFGDFW	1440	
1441	RMIWQRSATVMMTKLEERSVKCDQWPSRGTEHGLVQVTLDDTVELATYCVRTFAL	1500	
1441	RMIWQRSATVMMTKLEERSVKCDQWPSRGTEHGLVQVTLDDTVELATYCVRTFAL	1500	
1501	YKNGSEKREVRQFQFTAMPDHGVPEHPTPFLAFLLRRVKTCPNPPDAGPMVHVCSAGVGR	1560	
1501	YKNGSEKREVRQFQFTAMPDHGVPEHPTPFLAFLLRRVKTCPNPPDAGPMVHVCSAGVGR	1560	
1561	GCFTVIDAMLERIKHEKTVDIYGHVTLMAQRYNVQTEDQYIFHDALLEAVTCGTEV	1620	
1561	GCFTVIDAMLERIKHEKTVDIYGHVTLMAQRYNVQTEDQYIFHDALLEAVTCGTEV	1620	
1621	PARNLYAVIYOKLTQIETGENTVGMELPKRLASSKAHSTRFISANLPCNKFKNRLVINMP	1680	
1621	PARNLYAVIYOKLTQIETGENTVGMELPKRLASSKAHSTRFISANLPCNKFKNRLVINMP	1680	
1681	YESTRVCQLPIRGVSGSDYINASFIDGVRQKQKAYIATOGPLAETTEDFWRMLWEHNSSTIV	1740	
1681	YESTRVCQLPIRGVSGSDYINASFIDGVRQKQKAYIATOGPLAETTEDFWRMLWEHNSSTIV	1740	
1741	VMLTKLREMGREKCHOYPAERSARYQYFVDPMAEYNNPQVILREFKVTDARDGSRV	1800	
1741	VMLTKLREMGREKCHOYPAERSARYQYFVDPMAEYNNPQVILREFKVTDARDGSRV	1800	
1801	ROFQFTDPEQCVKPSGEGFIDFIOGVHKTKEQFGODGPISVHCSAGVGRGTGVFTLSIV	1860	
1801	ROFQFTDPEQCVKPSGEGFIDFIOGVHKTKEQFGODGPISVHCSAGVGRGTGVFTLSIV	1860	
1861	LERMEYEGVVDIFQTVKMLRTORPAMVQTEDQYQFSRAALEYLGFSFHYAT	1912	

1048	TSVLLSWEFPDNNYSPYKIQYNG--LTLVDGORTTKLITHLKPHTFNFVLTNRGSS	11048
1092	AGGLQHRVTAKTAPDVLRTKPAFIGTKNLDMITVOLPEVPANENIKGYIIIIIVPLKKSR	11514
1106	LGGLQQTVAWTA FNLLNGKPSVAPKDPADGFIWVYLPDGGSPVPVQSYFIVMWPLKRS	11650
1152	-GKFIKWSPDENMELDELKBEISR-KRRSIRYGREVEL-KPYIAAHFDVLPTEFTLGDD	12086
1166	GGQPLTPLGSPEDMDLBELIQDISRLQRRTVRHSRQLEVPRPYIAARFSVLPPTFHPGDQ	12222
1209	KHYGGFTNKOLQSGQEVYFVLA VMEHAESKMVATSPYSVPVVSMDLDPDPIITDEERGLI	12688
1226	KQYGGFDRGLPEGRHVRVLFVLA VJQKSE-PTFAASFSPDFFLDNDPQPIVDGEEGLI	12844
1269	WVYGPVLAVFIICIVIAILLYKR- ---KRAESDSRKSSIPNNKEI PSHTPTDPVELRRL	13244
1285	WVIGPVLAVFIICIVIAILLYKKNKPSKRKSEPRTKCLLNNADLAPHHPKQPVEMRI	13444
1325	NFQTP-----GMASHPPPIPILEADHIERLKANDNLIKFSQEVESIDPGQ	13680
1345	NFQTPDGLRSLKPEGFHPFESMLSHPPIPTADMAEHTERLKANDSLKLSQEVESIDPGQ	14044
1369	QFTWEHNLVNPKPNKYANVIAVDHRSVLLSATEGIPGSDYNNYVINDGYKQNAVIAT	14280
1405	QFTWEHNLVNPKPNKYANVIAVDHPRVILQPLEGIMGSDYINANYVDGYRRQNAYIAT	14644
1429	QGSLPETFGFWRMIWEQRSATVMMTKLBERSRVKCDQWPSRGTEHGLVQVTLDDTV	14880
1465	QGPLEPFGFWRNVWEQRSATVMMTRLEKSRIKCDQWPNRGITYFYGIQVTLDDTI	15244
1489	ELATYCVRTFALYKNGSSEKREVRQFQTAMPDGHVPEHPTPFLAFLRVRKTCNPPDAGP	15480
1525	ELATFCVRTFSLHKNGSSEKREVRQFQTAMPDGHVPEYPTPFLAFLRVRKTCNPPDAGP	15844
1549	MVYHCSAGVGRGCFIVIDAMLERI KHEKTVDIYGHVTLMRAQBNYVMQTEDQYIFHDA	16080
1585	IVYHCSAGVGRGCFIVIDAMLERI KPEKTVDVYGHVTLMRSQRNYVMQTEDQYSFHEA	16444
1609	LLEAVTCGNTVEPARNLYAYIQKLQJLTGBNTGMELEFKRLASSKAHTSRPISANLPC	16680
1645	LLEAVGCGNTEVPARSLYAYIQKLAQVEPGEHVTGMELEFKRLANSKAHTSRPISANLPC	17044
1669	NKFKRNLVNMPIYESTRVCLQPIRGVEGSDYINASFIDGYRQQKAYIATQGPLAETTEDF	17280
1705	NKFKRNLVNMPIYESTRVCLQPIRGVEGSDYINASFIDGYRQQKAYIATQGPLAETTEDF	17644
1729	WRMLWEHNSITVMLTKLRMGREKCHQYPAERSARYQYFVDPMAEYNNPQVILREFK	17880
1765	WRMLWENNSTIVMLTKLRMGREKCHQYPAERSARYQYFVDPMAEYNNPQVILREFK	18244
1789	VTDARDGQSRTRVQFQFTDMPQGVPKSGEGFIDFIGVHKTKFQFGDGPISVHCSAGV	18480
1825	VTDARDGQSRTRVQFQFTDMPQGVPKSGEGFIDFIGVHKTKFQFGDGPISVHCSAGV	18844
1849	GRTGVFTLSIVLMERYEGVVDIFQTVKMLRTQRPAMVQTEDQYQPSRYAALEYLGSFD	19080
1885	GRTGVFTLSIVLMERYEGVVDIFQTVKMLRTQRPAMVQTEDQYQCYQAALLEYLGSFD	19444
1909	HYAT 1912	
1945	HYAT 1948	

Qy	20	AETPPRTRTPVDGTGSSGVASFCIQTGDPBKIVNNKGGKKVSNORFEVIEDDGS	79
		: : : : : : : : : : : : : : : : :	
Dd	29	AEEPRIKEPKKOIGYSGSVASFVCATGDKPRTVNKGGKKVSORFETIEFDSEAG	88
		: : : : : : : : : : : : : : : : :	
Qy	80	SVLRIQLPRTRDRAIYECVASNVGEISVSTRITLVLRDDQIPRGFPPTDMGPOLKVVER	139
		: : : : : : : : : : : : : : : : :	
Dd	89	AVLRIQLPRTRDENVECVASQNSGEITVHAKITLVLRDDQIPSGFPNIDMGPOLKVVER	148
		: : : : : : : : : : : : : : : : :	
Qy	140	TRTATMLCAASGNPDPIITWFKPLVPDTNNNGRIKQLRSISEGGTPIRGALQITEQSEE	199
		: : : : : : : : : : : : : : : : :	

RESULT 3	
AEAD57121	
ID	AEAD57121 standard; protein; 1949 AA.
XX	
AC	AEAD57121;
XX	
DT	29-JAN-2004 (first entry)
XX	
DE	Human Protein U35234, SEQ ID NO 2981
XX	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2005, 19:37:10 ; Search time 52 Seconds
(without alignments)
2744.789 Million cell updates/sec

Title: US-10-772-636-64

Perfect score: 10042

Sequence: 1 MWHVRLLLLTFFLRTDA.....YQFSYRAALEYLGSPDHYAT 1912

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/1/iaa/5A COMB pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B COMB pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A COMB pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B COMB pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS COMB pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7245	72.1	1911	1	US-08-348-006B-5
2	7245	72.1	1911	1	Sequence 5, Appli
3	7245	72.1	1911	3	Sequence 5, Appli
4	7245	72.1	1911	5	Sequence 5, Appli
5	5745.5	57.2	1501	2	US-08-447-464-3
6	5745.5	57.2	1501	2	US-08-716-679-3
7	2646	26.3	538	4	US-09-743-492A-9
8	1587	15.8	793	1	US-08-015-985-3
9	1587	15.8	793	1	US-09-280-597-3
10	1565.5	15.6	802	1	US-08-015-985-1
11	1565.5	15.6	802	4	US-09-280-597-1
12	1565.5	15.6	807	4	US-09-949-016-7356
13	1542	15.4	1452	2	US-08-449-644-8
14	1542	15.4	1452	2	US-08-087-244A-8
15	1541	15.3	1452	2	US-08-652-971-4
16	1541	15.3	1452	2	US-08-991-258A-4
17	1541	15.3	1452	2	US-08-769-399-4
18	1541	15.3	1452	3	US-08-991-953A-4
19	1502.5	15.0	1439	2	US-08-449-644-2
20	1502.5	15.0	1439	2	US-08-087-244A-2
21	1482.5	14.8	1457	2	US-08-652-971-3
22	1482.5	14.8	1457	2	US-08-449-644-1
23	1482.5	14.8	1457	2	US-08-087-244A-1
24	1482.5	14.8	1457	2	US-08-991-258A-3
25	1482.5	14.8	1457	2	US-08-769-399-3
26	1482.5	14.8	1457	3	US-08-991-953A-3
27	1455	14.5	699	1	US-08-348-006B-7

28 1455 14.5 699 2 US-08-800-825A-7 Sequence 7, Appli
29 1455 14.5 699 3 US-09-158-657-7 Sequence 7, Appli
30 1450.5 14.4 1075 4 US-09-949-016-8308 Sequence 808, Ap
31 1415 14.1 296 1 US-10-374-539-3 Sequence 3, Appli
32 1384 13.8 289 1 US-08-036-210-13 Sequence 13, Appli
33 1384 13.8 289 2 US-08-449-609-13 Sequence 13, Appli
34 1384 13.8 289 4 US-09-361-096A-13 Sequence 13, Appli
35 1381.5 13.8 560 4 US-09-949-016-10786 Sequence 10786, A
36 1325 13.2 1436 2 US-08-652-971-2 Sequence 2, Appli
37 1325 13.2 1436 2 US-08-991-258A-2 Sequence 2, Appli
38 1325 13.2 1436 2 US-08-769-399-2 Sequence 2, Appli
39 1325 13.2 1436 3 US-08-991-953A-2 Sequence 2, Appli
40 1291.5 12.9 1445 1 US-08-015-986A-2 Sequence 2, Appli
41 1291.5 12.9 1445 2 US-08-446-363-2 Sequence 2, Appli
42 1286 12.8 1442 1 US-08-015-986A-3 Sequence 3, Appli
43 1286 12.8 1442 2 US-08-446-363-3 Sequence 26, Appli
44 1250 12.4 245 2 US-08-685-992-26 Sequence 26, Appli
45 1250 12.4 245 2 US-09-144-925-26 Sequence 26, Appli

ALIGNMENTS

RESULT 1

US-08-348-006B-5
; Sequence 5, Application US/08348006B
; Patent No. 5658756
; GENERAL INFORMATION:
; APPLICANT: RODAN, GIDEON A.
; APPLICANT: SCHMIDT, AZRIEL
; APPLICANT: RUTLEDGE, SU JANE
; TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN
; TITLE OF INVENTION: TYROSINE PHOSPHATASE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: J. MARK HAND
; STREET: 126 E. LINCOLN AVE., P.O. BOX 2000
; CITY: RAHWAY
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/348,006B
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/122,032
; FILING DATE: 14-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: HAND, J., MARK
; REGISTRATION NUMBER: 36,545
; REFERENCE/DOCKET NUMBER: 189921A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-594-3905
; TELEFAX: 908-594-4720
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1911 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-348-006B-5

Query Match 72.1%; Score 7245; DB 1; Length 1911;
Best Local Similarity 71.7%; Pred. No. 0;
Matches 1369; Conservative 219; Mismatches 279; Indels 42; Gaps 14;

QY 20 AETPPRFRTPVDTQVSGGVASFIQATGDRPKIWMNKKKKVGNQRFVEIEFDGSG 79
Db 29 AEEPFRFIKEPKDOIYSGRVASFVCOATGDPKPRVTWKKKKVNSQRFETIEFDESAG 88
QY 80 SVLRIOPLRTPRDEAIYECVASNNVGEISVSTRLTVLREDQIPRGPTTIDMGQPKLVKR 139
Db 89 AVLRIQPLRTPRDENVYECVAQNSVGEITVHAKLTVLREDQLPSGFNIDMGQPKLVKR 148
QY 140 TRTATMLCAASGNPDPEITWFKDPLVDTSNNGRIPKQLRSEISIGTPIRGALQIESSE 199
Db 149 TRTATMLCAASGNPDPEITWFKDPLVDPFSASNGRIKQLRSEISIGTPIRGALQIESSE 199
QY 200 SDGKTECVATNSAGTRYAPANLYVRELREVRVRPREFSIPTTHHEIMPGGSVNITCVA 259
Db 200 TDGKTECVATNSAGTRYSPANLYR---VRRVAPRFSILPMSHEIMPGGSVNITCVA 255
QY 260 VGSMPYVVKMLGAEDLTPEDDMPIGRNVLELNDVROQSAN---TCVAMSTLGVEATAQIT 318
Db 256 VGSMPYVVKMGAEDLTPEDDMPIGRNVLELNDVROQSAN---TCVAMSTLGVEATAQIT 315
QY 319 VKALPKPGTPVVTSTATSIITLWDSGNPEPVSYIIQHKPNSEBELYKEIDGVATTRY 378
Db 316 VKSLPKAPGPMVTENTATSIITLWDSGNPDVSYIIEYKSKSQDGPVOIKEDDITTRY 375
QY 379 SVAGLSYSDYPRVAVANNIGRPSEPVLITQTSQAPSSAPRDVQARMLSSTLILVOW 438
Db 376 SIGLSPNSEYEIWSAVNSIGOGPSESVVTRTGEQAPARPPRNQARMLSATMIVOW 435
QY 439 KEPEENGOIGYRVYVYTMPTQHVNNMKHNVADSOITIGNLVQKTVSVKVLAFSTI 498
Db 436 EEPVEPNGLIRGVYVYTMPEHPVGNWQHNDVDSLLTGVSLLEDETVTVRLAFTSV 495
QY 499 GDGLPSLDIOVITQTVGPOPLNFKAEPESETSILLSWTPPRSDTTANVELYVKDGEHGE 558
Db 496 GDGLPSLDIOVITQTVGPOPLNFKAEPESETSILLSWTPPRSDTTANVELYVKDGEHGE 558
QY 559 EORITTEPGTSYRLQGLKNSLYYFRLAASPOGLGASTAEISARTWQKSPSAPPQDISC 618
Db 556 EVGRTPDPTSYVVEDLKPNTYAFRLAASPOGLGASTAEISARTWQKSPSAPPQDISC 615
QY 619 TSPSSSTILVSWQPPVVEKONGIITEYSIKYTAVDGEDDKPHILGI PSDTTKYILLEOLE 678
Db 616 VSVRSTAILVSWRPPPETTINGALVGSVRYRPLGSEDPPEKVENGIPTTTOILLEALE 675
QY 679 KWTEYRITVTAHTDVGPPESLSVLIRTNEDVSPGPRKVEEAVNSTSVKUSWRSVPVN 738
Db 676 KWTOYRITVTAHTDVGPPESLSVLIRTNEDVSPGPRKVEEAVNSTSVKUSWRSVPVN 735
QY 739 KQHGIRGYQVHYVRMENGEPKQPMKQVMLADAQWFEFDDTTTEHDMITISGLQFETSYSL 798
Db 736 RQHGIRGYQVHYVRMENGEPKQPMKQVMLADAQWFEFDDTTTEHDMITISGLQFETSYSL 786
QY 799 TVTAYTTKGGGASKPKLVSTTGAVPKPKPLVNLNHTQMNTALLQWHPVVDFTF-GPIQGYR 857
Db 787 TVAAATYTKGGGASKPKLVSTTGAVPKPKPLVNLNHTQMNTALLQWHPVVDFTF-GPIQGYR 846
QY 858 LKFGKDMPELTLTLESEKEDHFTATDIIHKGASVFRLSARNKVGCEEMVKEISIPPEV 917
Db 847 LQFGREDSTPLATEFPSPSDRVTASGVHKGATVFRFLAARSFGGLGEEAAEVLISIPEDT 906
QY 918 PTGFPQMLHSEGTST-SVOLSWQPPVLAERNIGIITKYTLRYRDIINIPLLPMBQLIVPAD 976
Db 907 PRGHPQILEAGNASAGTVLLRWLPVPAERNGAIVKYTVAVAREAG-ALGPARETELPAG 965
QY 977 -----TTMTLTLGLKPDPTDVQKRAHSGPGPSVQFRLPVDQVPAKQHFHAKVM 1030
Db 966 RLSRARRTLLQGLKPDPTDVQKRAHSGPGPSVQFRLPVDQVPAKQHFHAKVM 1025
QY 1031 KTSVLLSWEPDNVNSAMPKILYDDGKMVEEDVGRATOKLIVNLKPEKSYSVFLNTRGN 1090
Db 1026 KTSVLLSWEPDNVNSAMPKILYDDGKMVEEDVGRATOKLIVNLKPEKSYSVFLNTRGN 1083
QY 1091 SAGGLQHRVTAKTAPDVLRTKPAFIGKTNLDGMITVOLPEVPANENIKGYIIIVPLKKS 1150

Db 1084 SLUGLQOQTVTAMTAFNLLNGKPSVAPKPDADGFIWVLPDQSPVPQSVFYIWMVPLRKS 1143
QY 1151 R-GKTIKPWESDEMEDELLEKEISR-KRSTRYGREVEL-KPYIAAHFDVLPTEFTLGD 1207
Db 1144 RGOQFLTPLGSPEDMDLEELIQDISRLQRRSLRHSRQLEVPRIYAARSVLPPTPHPGD 1203
QY 1208 DXHYGFTNKQJQSGOEYVFFVLAVMEHAESKMYATSPYSDPVSMDLDPQPIITDEEGL 1267
Db 1204 QKQYGGFNRLGEPGHRVYVFLVLAQKSE-TTFAASPSDFQLDNPDQPIVDEEGL 1262
QY 1268 INVGPVLAVFIIICIVTAILYKLR----KRAESRSKSSIPNNKBIPOSHPTDPVELRR 1323
Db 1263 IWWIGPVLAVFIIICIVTAILYKKNPKDSKRSKSDSPRTKCLLNADLAPHHKPDVEMRR 1322
QY 1324 LNPOTPGWASHPPPIPLELADHIEBLKANDNLKFSOYESIDPGQOFTWEHNSLVNKPK 1383
Db 1323 INFOTPGMLSHPPPIPIADWAEHTERLKANDSLKLSOYESIDPGQOFTWEHNSLVNKPK 1382
QY 1384 NRYANVIADHSRVLLSATEGIPGSDYNNANIDYGRKQNAVIAQGSLEPETFGDFWRMI 1443
Db 1383 NRYANVIADHSRVLLOPIEGIMSGSDYINANVYDGRQNAVIAQGSLEPETFGDFWRMI 1442
QY 1444 WQORSATVVMYTKLEERSVKCDQWPSRGTEHGLVQVTLTDLTVELATYCVRTPALYKN 1503
Db 1443 WQORSATVVMYTKLEERSVKCDQWPSRGTEHGLVQVTLTDLTVELATYCVRTPALYKN 1502
QY 1504 GSSEKREVRQOFTWPDHGVPEHTPELAFRRVKTCTNPPDAGPMVHCSAGVGTGCF 1563
Db 1503 GSSEKREVRQOFTWPDHGVPEHTPELAFRRVKTCTNPPDAGPMVHCSAGVGTGCF 1562
QY 1564 IVIDAMLERIKHEKTVDYIGHVTLMAQRYNVMVQTEQYIFTHDALLEAVTCGTEVVAR 1623
Db 1563 IVIDAMLERIKHEKTVDYIGHVTLMAQRYNVMVQTEQYIFTHDALLEAVTCGTEVVAR 1622
QY 1624 NUYAYIKLUTQETGENTGMELEFKRLASSKAHTSRFISANLPCNKKFQRLVNMIPYES 1683
Db 1623 SLVAYITQLAQVPEGHEVHTGMELEFKRLASSKAHTSRFISANLPCNKKFQRLVNMIPYES 1682
QY 1684 TRVCLQPIRGVSGDYINASFIDGRYQKQKAYATQGLAETTEDFWRMLWEHNSLIVVWL 1743
Db 1683 TRVCLQPIRGVSGDYINASFIDGRYQKQKAYATQGLAETTEDFWRMLWEHNSLIVVWL 1742
QY 1744 TKLREMGKREKCHQYPAERSARYQVFPVDPMAEYNNPQVILREFKVTDARDGQSRVQF 1803
Db 1743 TKLREMGKREKCHQYPAERSARYQVFPVDPMAEYNNPQVILREFKVTDARDGQSRVQF 1802
QY 1804 QFTDPEQGVKPSGEGFIDFIOVHKTKEQODGPISVHCSAGVGTGVTLSIVLER 1863
Db 1803 QFTDPEQGVKPSGEGFIDFIOVHKTKEQODGPISVHCSAGVGTGVTLSIVLER 1862
QY 1864 MRYEGVVDIFQTVKMLRTORPAMVQTEQYQFSYRAALEYLGSDHYAT 1912
Db 1863 MRYEGVVDIFQTVKMLRTORPAMVQTEQYQFSYRAALEYLGSDHYAT 1911

RESULT 2
US-08-800-825A-5
; Sequence 5, Application US/08800825A
; Patent No. 586397
; GENERAL INFORMATION:
; APPLICANT: RODAN, GIDEON A.
; APPLICANT: SCHMIDT, AZRIEL
; APPLICANT: RUTLEDGE, SU JANE
; TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN
; TITLE OF INVENTION: TYROSINE PHOSPHATASE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: J. MARK HAND - MERCK & CO., INC.
; STREET: 126 E. LINCOLN AVE., P.O. BOX 2000
; CITY: RAHWAY
; STATE: NEW JERSEY
; COUNTRY: USA

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2005, 19:41:01 ; Search time 220 Seconds
(without alignments)
3352.912 Million cell updates/sec

Title: US-10-772-636-64
Perfect score: 10042
Sequence: 1 MVHVARLLLLLLLLFFLRTDA.....YQSYRAALYLGSFDRHYAT 1912

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1722976 seqs, 385795295 residues

Total number of hits satisfying chosen parameters: 1722976

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*

2: /cgn2_6/ptodata/1/pubpaa/PCF_NEW_PUB.pep.*

3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*

4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*

5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*

6: /cgn2_6/ptodata/1/pubpaa/PCFUS_PUBCOMB.pep.*

7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*

8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*

9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*

10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*

11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*

12: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*

13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*

14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*

15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*

16: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*

17: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*

18: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*

19: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

20: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*

22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10042	100.0	1912	17	US-10-772-636-64
2	7369.5	73.4	1948	9	US-09-808-602-55
3	7369.5	73.4	1948	10	US-09-800-198-45
4	7037	70.1	1907	15	US-10-291-265-250
5	7032	70.0	1897	17	US-10-482-023-52
6	7032	70.0	1897	18	US-10-712-892A-34
7	5794.5	57.7	1502	9	US-09-808-602-54
8	5794.5	57.7	1502	10	US-09-800-198-44
9	5559.5	55.4	1495	15	US-10-258-666-12
10	4877	48.6	2029	15	US-10-087-684-38
11	4877	48.6	2029	15	US-10-218-779-39

12	4865	48.4	2037	15	US-10-087-684-39	Sequence 39, Appl
13	4865	48.4	2037	15	US-10-218-779-39	Sequence 39, Appl
14	3027.5	30.1	647	15	US-10-291-265-722	Sequence 722, App
15	2965.5	29.5	1194	14	US-10-191-029-10	Sequence 10, Appl
16	2238	22.3	1585	15	US-10-369-493-5569	Sequence 5569, Ap
17	2238	22.3	1585	15	US-10-369-493-5570	Sequence 5570, Ap
18	2166	21.6	442	9	US-09-925-300-950	Sequence 950, App
19	1814	18.1	344	16	US-10-408-765A-1670	Sequence 1670, Ap
20	1587	15.8	793	17	US-10-777-144-3	Sequence 3, Appli
21	1587	15.8	793	17	US-10-777-145-3	Sequence 3, Appli
22	1587	15.8	793	17	US-10-777-186-3	Sequence 3, Appli
23	1565.5	15.6	802	17	US-10-777-144-1	Sequence 1, Appli
24	1565.5	15.6	802	17	US-10-777-145-1	Sequence 1, Appli
25	1565.5	15.6	807	17	US-10-777-186-1	Sequence 1, Appli
26	1565.5	15.6	807	15	US-10-264-049-3138	Sequence 3138, Ap
27	1548	15.4	1452	9	US-09-887-669-8	Sequence 8, Appli
28	1529	15.2	1452	16	US-10-408-765A-83	Sequence 83, Appl
29	1505.5	15.0	1439	9	US-09-887-669-2	Sequence 2, Appli
30	1502.5	15.0	1439	16	US-10-408-765A-284	Sequence 284, App
31	1482.5	14.8	1457	9	US-09-887-669-1	Sequence 1, Appli
32	1461	14.5	774	13	US-10-087-192-1017	Sequence 1017, Ap
33	1460.5	14.5	1444	15	US-10-058-270A-98	Sequence 98, Appl
34	1454	14.5	700	15	US-10-366-547-63	Sequence 63, Appl
35	1454	14.5	700	15	US-10-444-795B-793	Sequence 793, App
36	1454	14.5	700	16	US-10-473-127-912	Sequence 912, App
37	1454	14.5	700	16	US-10-473-127-915	Sequence 915, App
38	1454	14.5	700	16	US-10-473-127-916	Sequence 916, App
39	1454	14.5	700	16	US-10-473-127-919	Sequence 919, App
40	1454	14.5	700	16	US-10-473-127-918	Sequence 918, App
41	1454	14.5	717	16	US-10-753-267-100	Sequence 100, App
42	1454	14.5	717	16	US-10-473-127-917	Sequence 917, App
43	1454	14.5	1463	14	US-10-176-847-22	Sequence 22, Appl
44	1454	14.5	1463	14	US-10-205-823-343	Sequence 343, App
45	1444	14.4	699	15	US-10-366-547-65	Sequence 65, Appl

ALIGNMENTS

RESULT 1

US-10-772-636-64

; Sequence 64, Application US/10772636

; Publication No. US20050042687A1

; GENERAL INFORMATION:

; APPLICANT: Kelly, Louise M.

; APPLICANT: Carroll, Joseph M.

; APPLICANT: Farlow, Deborah

; APPLICANT: Healy, Aileen

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING

; TITLE OF INVENTION: HEMATOLOGICAL DISORDERS USING 9118, 990, 17662, 81982, 630,

; TITLE OF INVENTION: 21472, 17692, 19290, 21620, 21689, 28899, 53659, 64549,

; TITLE OF INVENTION: 9465, 23544, 7366, 27417, 57259, 21844, 943, 2061, 5891,

; TITLE OF INVENTION: 9137, 13908, 14310, 17600, 25584, 27824, 28469, 38947,

; TITLE OF INVENTION: 53003, 965, 56639, 9661, 16052, 1521, 6662, 13913, 12405 OR

; FILE REFERENCE: MP103-015PIRNONMIM

; CURRENT APPLICATION NUMBER: US/10/772,636

; CURRENT FILING DATE: 2004-02-05

; PRIOR APPLICATION NUMBER: US 60/445,241

; PRIOR FILING DATE: 2003-02-05

; PRIOR APPLICATION NUMBER: US 60/448,389

; PRIOR FILING DATE: 2003-02-18

; PRIOR APPLICATION NUMBER: US 60/456,320

; PRIOR FILING DATE: 2003-03-20

; PRIOR APPLICATION NUMBER: US 60/460,279

; PRIOR FILING DATE: 2003-04-03

; PRIOR APPLICATION NUMBER: US 60/465,924

; PRIOR FILING DATE: 2003-04-28

; PRIOR APPLICATION NUMBER: US 60/470,052

; PRIOR FILING DATE: 2003-05-13

; PRIOR APPLICATION NUMBER: US 60/498,106

; PRIOR FILING DATE: 2003-08-26

; PRIOR APPLICATION NUMBER: US 60/500,179

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run On: July 1, 2005, 19:35:47 ; Search time 61 Seconds
(without alignments)
3015.843 Million cell updates/sec

Title: US-10-772-636-64
Perfect score: 10042
Sequence: 1 MVHVARLLLLLLTFLRTDA.....YQFSYRAALEYLGSFDHYAT 1912

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10042	100.0	1912	2	A56178 protein-tyrosine-p
2	9433.5	93.9	1894	2	C54689 protein-tyrosine-p
3	8394.5	83.6	1691	1	D54689 protein-tyrosine-p
4	7117.5	70.9	1907	2	S50893 protein-tyrosine-p
5	7032	70.0	1897	1	TDHULK leukocyte antigen-
6	6981.5	69.5	1898	2	S46216 leukocyte antigen-
7	6971.5	69.4	1863	2	S46217 protein-tyrosine-p
8	5871.5	58.5	1499	2	I50212 protein-tyrosine-p
9	5745.5	57.2	1501	2	I58148 protein-tyrosine-p
10	5562	55.4	1496	1	A48758 protein-tyrosine-p
11	4877	48.6	2029	1	TDPELK protein-tyrosine-p
12	4650.5	46.3	1290	2	A56493 leucocyte common a
13	4315.5	43.0	1262	1	B48758 protein-tyrosine-p
14	4127	41.1	2051	2	T30938 receptor tyrosine
15	3372	33.6	1231	2	T31093 probable protein-t
16	2996	29.8	1237	2	S53089 protein-tyrosine-p
17	2820	28.1	582	2	A57068 protein-tyrosine-p
18	2238	22.3	1585	2	T19121 probable protein-t
19	1954	19.5	398	2	I56540 protein-tyrosine-p
20	1565.5	15.6	802	1	A36065 protein-tyrosine-p
21	1553	15.5	829	1	A47373 protein-tyrosine-p
22	1542	15.4	1452	1	S17669 protein-tyrosine-p
23	1541	15.3	1452	1	S17670 protein-tyrosine-p
24	1539.5	15.3	796	1	JCL285 protein-tyrosine-p
25	1539	15.3	832	2	JC8051 protein tyrosine p
26	1510	15.0	440	2	I50213 protein-tyrosine-p
27	1486	14.8	1440	2	JC6312 protein-tyrosine-p
28	1482.5	14.8	1457	1	A48066 protein-tyrosine-p
29	1455	14.5	699	2	JC6132 protein-tyrosine-p

30	1454	14.5	700	1	S12053 protein-tyrosine-p
31	1372	13.7	680	2	JC8052 protein-tyrosine-p
32	1331.5	13.3	1436	2	JC5290 protein-tyrosine-p
33	1286	12.8	1442	1	B48148 protein-tyrosine-p
34	1285.5	12.8	1445	1	A48148 protein-tyrosine-p
35	1270.5	12.7	1442	2	S72441 protein-tyrosine-p
36	1266	12.6	1422	2	T42636 protein-tyrosine-p
37	1242	12.4	1301	1	A41622 protein-tyrosine-p
38	1226.5	12.2	2302	2	T14328 protein-tyrosine-p
39	1205.5	12.0	1462	1	B36182 protein-tyrosine-p
40	1204	12.0	2314	1	A46151 protein-tyrosine-p
41	1100	11.0	1273	1	TDRTLT leukocyte common a
42	1097	10.9	818	2	T19120 hypothetical prote
43	1088	10.8	1237	2	A54080 protein-tyrosine-p
44	1060	10.6	1304	1	A46546 leucocyte common a
45	1056	10.5	1291	1	A28334 protein-tyrosine-p

ALIGNMENTS

RESULT 1

A56178
N:Alternate names: protein-tyrosine-phosphatase BPTP-2
C:Species: Homo sapiens (man)
C:Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 09-Jul-2004
C:Accession: A56178; S12052; B4929
R:Pulido, R.; Krueger, N.X.; Serra-Pages, C.; Saito, H.; Streuli, M.
J. Biol. Chem. 270, 6722-6728, 1995
A:Title: Molecular characterization of the human transmembrane protein-tyrosine phosphatase delta isoforms.
A:Reference number: A56178; MUID:95204468; PMID:7896816
A:Accession: A56178
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1912 <PUL>
A:Cross-references: UNIPROT:P23468; GB:L38929; NID:g755652; PIDN:AAC41749.1; PID:g755653
R:Krueger, N.X.; Streuli, M.; Saito, H.
EMBO J. 9, 3241-3252, 1990
A:Title: Structural diversity and evolution of human receptor-like protein tyrosine phosphatase delta isoforms.
A:Reference number: S12049; MUID:91006018; PMID:2170109
A:Accession: S12052
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 350-1912 <KRU>
A:Cross-references: GB:X54133; NID:g35789; PIDN:CAA38068.1; PID:g35790
A:Note: the sequence from Fig. 5B is inconsistent with that from Fig. 5A in having 568-T
R:Adachi, M.; Sekiya, M.; Azimura, Y.; Takekawa, M.; Itoh, F.; Hinoda, Y.; Imai, K.; Yac
Cancer Res. 52, 737-740, 1992
A:Title: Protein-tyrosine phosphatase expression in pre-B cell NALM-6.
A:Reference number: A44929; MUID:92119637; PMID:1370651
A:Accession: B44929
A:Molecule type: mRNA
A:Residues: 1756-1804, 'C', 1806-1845 <ADA>
A:Cross-references: GB:S78086; NID:g243545; PIDN:AAB21147.1; PID:g243546
A:Experimental source: pre-B cell NALM-6
A:Note: sequence extracted from NCBI backbone (NCBIN:78086, NCBIP:78087)
A:Note: the authors did not report the entire codon for residue 90
C:Genetics:
A:Gene: GDB:PTPRD
A:Cross-references: GDB:I131384; OMIM:601598
A:Map position: 9p24-9p24
C:Superfamily: leukocyte antigen-related protein; fibronectin type III repeat homology; ogy
C:Keywords: glycoprotein; phosphoprotein; phosphoric monoester hydrolase; transmembrane
F:38-100/Domain: immunoglobulin homology <IMM1>
F:140-209/Domain: immunoglobulin homology <IMM2>
F:250-304/Domain: immunoglobulin homology <IMM3>
F:711-811/Domain: fibronectin type III repeat homology <IMM4>
F:1293-1912/Domain: leukocyte common antigen cytosolic domain homology <3PR>
F:1669-1892/Domain: protein-tyrosine-phosphatase homology <PTP2>
F:1553/Active site: Cys (phosphocysteine intermediate) #status predicted

F,1559/Binding site: substrate phosphate (Arg) #status predicted									
F,1844/Active site: Cys (phosphocysteine intermediate) #status predicted									
F,1850/Binding site: substrate phosphate (Arg) #status predicted									
Query Match 100.0%; Score 10042; DB 2; Length 1912;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 1912; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	MVHVARLLLLLTFFLRDAETPRFTPTVDQTVSGGVASFICQATGDP	PRPKI	VNKK	60				
Db	1	MVHVARLLLLLTFFLRDAETPRFTPTVDQTVSGGVASFICQATGDP	PRPKI	VNKK	60				
Qy	61	GKVSQNRFEVIEFDDGSGSVLRIOQLRTPRDEAIYECVANNVGEISV	TRLTVL	REDQ	120				
Db	61	GKVSQNRFEVIEFDDGSGSVLRIOQLRTPRDEAIYECVANNVGEISV	TRLTVL	REDQ	120				
Qy	121	IPRGFPTIDMGPOLKVERTRTATMLCAASGNPDPEITWFKDPL	PVDT	SNNGRI	KQLRS	180			
Db	121	IPRGFPTIDMGPOLKVERTRTATMLCAASGNPDPEITWFKDPL	PVDT	SNNGRI	KQLRS	180			
Qy	181	ESIGTPIRGALQIEQSESDQGYECVATNSAGTRYSPANLVREL	REVR	PPRF	SI	240			
Db	181	ESIGTPIRGALQIEQSESDQGYECVATNSAGTRYSPANLVREL	REVR	PPRF	SI	240			
Qy	241	PPTNHEIMPGGSVNITCVAVGSPMPYVKWMLGAEDLTPEDDMP	IGRNV	LENDV	RQSAN	300			
Db	241	PPTNHEIMPGGSVNITCVAVGSPMPYVKWMLGAEDLTPEDDMP	IGRNV	LENDV	RQSAN	300			
Qy	301	TCVAMSTLGVTEAIAQITVKALPDPGTPVTESTATSI	TLTW	DSGN	PEPVSYII	IOHKP	360		
Db	301	TCVAMSTLGVTEAIAQITVKALPDPGTPVTESTATSI	TLTW	DSGN	PEPVSYII	IOHKP	360		
Qy	361	KNSLELYKEIDGVATTRYSVAGLSYSDYERFVAVANNIGRPP	SEPV	LIT	TSOAP	SSA	420		
Db	361	KNSLELYKEIDGVATTRYSVAGLSYSDYERFVAVANNIGRPP	SEPV	LIT	TSOAP	SSA	420		
Qy	421	PRDVOARMLSTTILVQWKEPEEPNGQIQGYRVYTTMDPTQ	OVANN	MMKHN	VAD	SOIT	480		
Db	421	PRDVOARMLSTTILVQWKEPEEPNGQIQGYRVYTTMDPTQ	OVANN	MMKHN	VAD	SOIT	480		
Qy	481	NLVPOKTVSVKVLAFSTSGDGPLASDIQVITQTGVPGQPLN	FKAP	EPSE	TSILL	SWTP	540		
Db	481	NLVPOKTVSVKVLAFSTSGDGPLASDIQVITQTGVPGQPLN	FKAP	EPSE	TSILL	SWTP	540		
Qy	541	SDDTANYELVYKDGSHGEQRITTEPGTSYRLQGLKPNLS	LYFRLA	ARS	POGL	GAST	600		
Db	541	SDDTANYELVYKDGSHGEQRITTEPGTSYRLQGLKPNLS	LYFRLA	ARS	POGL	GAST	600		
Qy	601	SARTWQSKPSAPPQDISCTSPSSSTILVSWQPPVPEKQNG	IITEYS	KYTA	VDGED	DKPH	660		
Db	601	SARTWQSKPSAPPQDISCTSPSSSTILVSWQPPVPEKQNG	IITEYS	KYTA	VDGED	DKPH	660		
Qy	661	BILGIPSTTKYLLEQLEKWTYRITVTAHTDVGPPESLS	VLIR	TNE	DVPS	GP	720		
Db	661	BILGIPSTTKYLLEQLEKWTYRITVTAHTDVGPPESLS	VLIR	TNE	DVPS	GP	720		
Qy	721	EAVNSTSVKWSRSPVNKQHQIRGYOVHYVRMENGEPKQ	PMKOV	ML	ADA	QW	780		
Db	721	EAVNSTSVKWSRSPVNKQHQIRGYOVHYVRMENGEPKQ	PMKOV	ML	ADA	QW	780		
Qy	781	TEHDMIIISGLQPEYSYSLTVTAYTTTKDGA	RKPKLV	STT	GAV	PKPL	840		
Db	781	TEHDMIIISGLQPEYSYSLTVTAYTTTKDGA	RKPKLV	STT	GAV	PKPL	840		
Qy	841	IQWHPVDTFGPOLQYRLKFKGKOMEPLTTLEFSEKEDH	FTAT	DI	HGAS	TV	900		
Db	841	IQWHPVDTFGPOLQYRLKFKGKOMEPLTTLEFSEKEDH	FTAT	DI	HGAS	TV	900		
Qy	901	VGFGEEMKEISIPREVTPGPQNLSHSGTSTSVLSQW	QPVLA	ERN	GI	IT	960		
Db	901	VGFGEEMKEISIPREVTPGPQNLSHSGTSTSVLSQW	QPVLA	ERN	GI	IT	960		
Qy	961	INIPLLPMEQLIVPADTTMTLTGLKPDPTTYDVK	VRAH	TS	KG	PG	1020		

Db	961	INIPLLPMEQLIVPADTTMTLTGLKPDPTTYDVK	VRAH	TS	KG	PG	1020		
Qy	1021	AKNFHVKAVMKTSVLLSWEIPENVNSAMPFKILYDDG	KMVEE	VD	GRAT	QKLI	1080		
Db	1021	AKNFHVKAVMKTSVLLSWEIPENVNSAMPFKILYDDG	KMVEE	VD	GRAT	QKLI	1080		
Qy	1081	YSFVLNTRNGNSAGGLQHRVTAKTAPDVLRTKPA	FIGK	TN	LD	GM	1140		
Db	1081	YSFVLNTRNGNSAGGLQHRVTAKTAPDVLRTKPA	FIGK	TN	LD	GM	1140		
Qy	1141	YIIIVPLKKSGRGFIKPWESPDEWELDELLKEIS	RKRS	IR	Y	GRE	1200		
Db	1141	YIIIVPLKKSGRGFIKPWESPDEWELDELLKEIS	RKRS	IR	Y	GRE	1200		
Qy	1201	TEFTLGDGDKHYGFTNKQLQSGQGYVFFVLA	VN	HAES	KMYAT	SP	1260		
Db	1201	TEFTLGDGDKHYGFTNKQLQSGQGYVFFVLA	VN	HAES	KMYAT	SP	1260		
Qy	1261	TDSEEGLIWVGVPLAVVFIICIVIAILLYK	KRAES	DS	RK	SS	1320		
Db	1261	TDSEEGLIWVGVPLAVVFIICIVIAILLYK	KRAES	DS	RK	SS	1320		
Qy	1321	LRRLNFOTPGMAHPPPIPILELADHIERL	KAND	NL	K	F	1380		
Db	1321	LRRLNFOTPGMAHPPPIPILELADHIERL	KAND	NL	K	F	1380		
Qy	1381	KPKNRVANVTAYDHSRVLLSAIEGIPGSD	VN	AN	Y	IG	1440		
Db	1381	KPKNRVANVTAYDHSRVLLSAIEGIPGSD	VN	AN	Y	IG	1440		
Qy	1441	RMIWEORSATVVMVMTKLEERSVKCDQ	WP	SR	G	T	1500		
Db	1441	RMIWEORSATVVMVMTKLEERSVKCDQ	WP	SR	G	T	1500		
Qy	1501	YKNGSSEKREVROQFTAMPDGHVP	PHPT	PLA	FL	RV	1560		
Db	1501	YKNGSSEKREVROQFTAMPDGHVP	PHPT	PLA	FL	RV	1560		
Qy	1561	GCFTVIDAMLERIKHEKTVDIYGHV	TL	MR	AO	RV	1620		
Db	1561	GCFTVIDAMLERIKHEKTVDIYGHV	TL	MR	AO	RV	1620		
Qy	1621	PARNLYAYIQKLQIETGENVTGMELE	FK	L	A	S	1680		
Db	1621	PARNLYAYIQKLQIETGENVTGMELE	FK	L	A	S	1680		
Qy	1681	YESTRVCLOPIRGVEGSDYINASF	ID	GY	R	O	1740		
Db	1681	YESTRVCLOPIRGVEGSDYINASF	ID	GY	R	O	1740		
Qy	1741	VMLTKLREMGREKCHQYMPAERSA	RYOY	F	V	D	1800		
Db	1741	VMLTKLREMGREKCHQYMPAERSA	RYOY	F	V	D	1800		
Qy	1801	RQFQFTDWPQGVKPGEGEFDIF	IG	Q	V	H	1860		
Db	1801	RQFQFTDWPQGVKPGEGEFDIF	IG	Q	V	H	1860		
Qy	1861	LERMRVEGVVDIQTVMKLTQR	PAM	V	Q	T	1912		
Db	1861	LERMRVEGVVDIQTVMKLTQR	PAM	V	Q	T	1912		

RESULT 2

C54689

protein-tyrosine-phosphatase (SC 3.1.3.48), receptor type delta, splice form B precursor
N;Alternate names: WPTP delta type B/C

N;Contains: protein tyrosine phosphatase, receptor type delta, splice form C

C;Species: Mus musculus (house mouse)

C;Date: 25-Apr-1995 #sequence_revision 19-May-1995 #text_change 09-Jul-2004

C;Accession: C54689; B54689

R;Mizuno, K.; Hasegawa, K.; Katagiri, T.; Ogimoto, M.; Ichikawa, T.; Yakura, H.

536	Db	EDTIASVELVDRDGDQGEORIIITBPGTSYRLQGLKPNLSLYYFRLSATSPQGLGASTAEI	595
601	Qy	SARTMOSKAPQDII SCTSPSSTISILVSWQPPPEVKONGIIITEYSIKYTAVDGEDDKPH	660
596	Db	SARTMQ-KPSAPQDII SCTSPSSTISILVSWQPPPEVKONGIIITEYSUKYAAVDGEDYKPH	654
661	Qy	EILGIPSDTTYKYLLEQLEKWTEYRIITVTAHTDVGPGPESLSVLIRTNEDVPSPRKRKEV	720
655	Db	EIIGNSDTTYKYLLEQLEKWTEYRIITVTAHTDVGPPWPESSLVIRTDDEVPSPRKRKEV	714
721	Qy	EAVNSTSVKVSWSPPNKHQGOIRGYOVHYVPMENGEPKQPMKDKVMLADAQWEDDT	780
715	Db	EAVNATAVKVSWSPNKHQGOIRGYOVHYVPMENGEPK-SAMLKDKVMLADAQ-----	767
781	Qy	TEHDMIISGLQPTSYSLTVATYTKGDGARSXPKLVSTTGAVGPKPRLVINHTOMTAL	840
768	Db	---DMIISGLQPTSYSLTVATYTKGDGARSXPKLVSTTGAVGPKPRLVINHTOMTAL	824
841	Qy	IOWHPPVDTFGLQGYRLKFGRKDMBPLTTLFSEKEDHFTATDIIHKGASVYFRLSARNK	900
825	Db	IOWHPPVDTFG-LQGYRLKFGRKDMBPLTTLFSEKEDHFTATDIIHKGASVYFRLSARNK	883
901	Qy	VGGEEMVXKEISIPREVPFGFPONLHSEGTSTSVQLSWQPPVLAERNGIITKYTLLYRD	960
884	Db	VGGEEMVXKEISVPEBIPTGFPONLHSEGTSTSVQLSWQPPVLAERNGIITKYTLLYRD	943
961	Qy	INIPLLPMBEOLIPADTWTMLTGLKPDTTYDVKVRHAHTSKGPGYSPSVQFRTL PVDO-V	1019
944	Db	INIPLLPMBEHLIPADTSMTLTGLKSDTTYDVKVRHAHTSKRGPGYSPSVQFRTL PVDOQM	1003
1020	Qy	FAKNPHVKAVMKTSVLLSWEIIPENYNSAMPFK-ILYDDGGMVBEVDGRATQKLI VNLKPE	1078
1004	Db	FAKNPHVKAVMKTSVLLSWEIIPENYNPA ILSKFFMMMDGMVBEVDGRATQKLI VNLKPE	1063
1079	Qy	KSYSFVLVTRNGSAGGLQHRVTAKTADVLRTKPAFIGKTNLDGMTVQLPEVPANENIK	1138
1064	Db	KSYSFVLVTRNGSAGGLQHRVTAKTADVLRTKPAFIGKTNLDGMTVQLPDVPANENIK	1123
1139	Qy	GYVIIIVPLKKSXGKFIKWESPDEMEDELLEKEISRKRRSIRYGREVELKPYIAAHFDV	1198
1124	Db	GYVIIIVPLKKSXGKFIKWESPDEMEDELLEKEISRKRRSIRYGREVELKPYIAAHFDV	1183
1199	Qy	LPTEFTLGDHXYGFTNKQLSQGGYVYFVLAVMHAESKMYATSPYSDPVVSMOLDPQ	1258
1184	Db	LPTEFTLGDHXYGFTNKQLSQGGYVYFVLAVMHAESKMYATSPYSDPVVSMOLDPQ	1243
1259	Qy	PITDEEGIIWVGPVLAVVFIICIVIAILLYKRGKAESDSRKSSIPNNKEIPSHUPTDP	1318
1244	Db	PITDEEGIIWVGPVLAVVFIICIVIAILLYKRGKAESDSRKSSIPNNKEIPSHUPTDP	1303
1319	Qy	VELURLNFOTPGWASHPPPILELADHIERLKANDNLKFSQYESIDPGOOFTWEHSNLE	1378
1304	Db	VELURLNFOTPGWASHPPPILELADHIERLKANDNLKFSQYESIDPGOOFTWEHSNLE	1363
1379	Qy	VNPKNRYANVIAYDHSRVLLSAIEGIPGSDYVNANYIDGIRKQNAVIAIATQGSLEPTFGD	1438
1364	Db	VNPKNRYANVIAYDHSRVLLSAIEGIPGSDYVNANYIDGIRKQNAVIAIATQGSLEPTFGD	1423
1439	Qy	FWRMIWEQASATVMMTKLEERSVKCDQYWPISRGTTETHGLVQVTLTLDVELATYCVRTF	1498
1424	Db	FWRMIWEQ-EATVMMTKLEERSVKCDQYWPISRGTTETHGLVQVTLTLDVEL-TYCVRTF	1481
1499	Qy	ALYKNGSSSEKREVRQFQTAWPDHGVPEHPTTFLAFLRRVKTCNPPDAGPMVHVCSAGVG	1558
1482	Db	ALYKNGSSSEKRVQRQFQTAWPDHGVPEHPTTFLAFLRRVKTCNPPDAGPMVHVCSAGVG	1541
1559	Qy	RTGCFIVIDLAMLERIKHEKTVDIYGHVTLWRAQRNVMVQTEQYIFIHDLALLEAVTCGNT	1618
1542	Db	RTGCFIVIDLAMLERIKHEKTVDIYGHVTLWRAQRNVMVQTEQYIFIHDLALLEAVTCGNT	1601
1619	Qy	EVPARNLYAIYOKLTOIETGENVTGMLEFKRLASKAHTSRFISANLPCNKFKRLVNI	1678
1602	Db	EVPARNLYAIYOKLTOIETGENVTGMLEFKRLASKAHTSRFISANLPCNKFKRLVNI	1661

Qy 1679 MPYESTRVCLOPIRGVSGDYINASFIDGVRQKQAYIATOGPLAETTFEDFRLWEHNST 1738
 Db 1662 MPYESTRVCLOPIRGVSGDYINASFIDGVRQKQAYIATOGPLAETTFEDFRLWEHNST 1721
 Qy 1739 IVNMLTKLRNGREKCHQYWPASARSARYQYFVDPMAEYNNPQVILREFKVTVDARDQQR 1798
 Db 1722 IVNMLTKLRNGREKCHQYWPASARSARYQYFVDPMAEYNNPQVILREFKVTVDARD-QSR 1780
 Qy 1799 TVRQFQFTDPEQGVKSGEGFIDFQGVHKTKEQFGODGPISVHCSAGVGTGVFTILS 1858
 Db 1781 TVRQFQFTDPEQGVKSGEGFIDFQGVHKTKEQFGODGPISVHCSAGVGTGVFTILS 1840
 Qy 1859 IVLERMYEGVVDIQTVMKLTORPAMVQTEQYQSYRAALSYLGSFQHYAT 1912
 Db 1841 IVLERMYEGVVDIQTVMKLTORPAMVQTEQYQSYRAALSYLGSFQHYAT 1894

RESULT 3
 D54689
 protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type delta, splice form D precursor
 N:Alternate names: MPTP delta type D
 N:Contains: protein tyrosine phosphatase, receptor type delta, splice form A
 C:Species: Mus musculus (house mouse)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C:Accession: D54689; A54689
 R:Miuno, K.; Hasegawa, K.; Katagiri, T.; Ogimoto, M.; Ichikawa, T.; Yakura, H.
 Mol. Cell. Biol. 13, 5513-5523, 1993
 A:Title: MPTP delta, a putative murine homolog of HPTP delta, is expressed in specialize
 A:Reference number: A54689; MUID:93360986; PMID:835697
 A:Accession: D54689
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1691 <M12>
 A:Cross-references: UNIPROT:Q64487
 A:Experimental source: brain
 A:Note: sequence inconsistent with nucleotide translation
 A:Note: sequence extracted from NCBI backbone (NCBIN:137486, NCBI:P:136537)
 A:Accession: A54689
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-398,799-1691 <M12>
 A:Experimental source: brain
 A:Note: sequence inconsistent with nucleotide translation
 A:Note: sequence extracted from NCBI backbone (NCBIN:136522, NCBI:P:136524)
 C:Superfamily: leukocyte antigen-related protein; fibronectin type III repeat homology;
 ogy
 C:Keywords: alternative splicing; glycoprotein; phosphoprotein; phosphoric monoester hyd
 F:42-95/Domain: immunoglobulin homology <IMM3>
 F:114-196/Domain: fibronectin type III repeat homology <FN3A>
 F:1075-1691/Domain: leukocyte common antigen cytosolic domain homology <LAC>
 F:1449-1671/Domain: protein-tyrosine-phosphatase homology <PTP2>
 F:1333/Active site: Cys (phosphocysteine intermediate) #status predicted
 F:1339/Binding site: substrate phosphate (Arg) #status predicted
 F:1623/Active site: Cys (phosphocysteine intermediate) #status predicted
 F:1629/Binding site: substrate phosphate (Arg) #status predicted

Query Match 83.6%; Score 8394.5; DB 1; Length 1691;
 Best Local Similarity 95.7%; Pred. No. 0;
 Matches 1611; Conservative 28; Mismatches 28; Indels 17; Gaps 9;

Qy 231 VRRVPPRESIPTNHEIMPGGSVNITCVAGSPMPYVKWMLGAEDLTPEDDMPITGRNVLE 290
 Db 23 VRRVPPRESIPTNHEIMPGGSVNITCVAGSPMPYVKWMLGAEDLTPEDDMPITGRNVLE 82
 Qy 291 LNDVRQSNYTCVAMSTLGVIEAIAQITVKALPKPPGTPVVTSTATSIITLWDSGNPEP 350
 Db 83 LNDVRQSNYTCVAMSTLGVIEAIAQITVKALPKPPGTPVVTSTATSIITLWDSGNPEP 142
 Qy 351 VSYIIOHKPNSEELYKEIDGVAATRYSVAGLSYSPYEFRRVAVNNIGRGPSEPVLT 410
 Db 143 VSYIIOHKPNSEELYKEIDGVAATRYSVAGLSYSPYEFRRVAVNNIGRGPSEPVLT 202

Qy 411 QTSEQAPSSAPRDVQARMLSSTIILVQWKEPEBPNGQIQGYRVYYTMDPTQHVNMMKHN 470
 Db 203 QTSEQAPSSAPRDVQARMLSSTIILVQWKEPEBPNGQIQGYRVYYTMDPTQHVNMMKHN 262
 Qy 471 VADSQITITGNLVQPKTYSVKVLAFTSISGDPSSDIQVITQTGVPGQPLNFKAEPSSET 530
 Db 263 VADSQITITGNLVQPKTYSVKVLAFTSISGDPSSDIQVITQTGVPGQPLNFKAEPSSET 322
 Qy 531 SILLSTPPRSDIANVELVYKDGHEGEORITIEPCTSRLQGLKPNLSLYFRLAARSP 590
 Db 323 SILLSTPPRSDIANVELVYKDGHEGEORITIEPCTSRLQGLKPNLSLYFRLAARSP 382
 Qy 591 QGLGASTAEISARTMOSKPSAPPQDISCTSPSSSTILVSWQPPPPVKKONGIIEYSIKYT 650
 Db 383 QGLGASTAEISARTMOSKPSAPPQDISCTSPSSSTILVSWQPPPPVKKONGIIEYSIKYT 441
 Qy 651 AVDGEDDKPHEILGIPSDTTKYLLEQLEKWTYRITVTAHTDVGPGPESLVLRTNEDV 710
 Db 442 AVDGEDDKPHEILGIPSDTTKYLLEQLEKWTYRITVTAHTDVGPGPESLVLRTNEDV 501
 Qy 711 PSQPPRKVEAVNSTSVKSWESPVNKHOGQIRGVQVHYVRMENGEPKQGMKLDVML 770
 Db 502 PSQPPRKVEAVNSTSVKSWESPVNKHOGQIRGVQVHYVRMENGEPK-SAMKLDVML 560
 Qy 771 ADAQWEFDDTTEHDMIIISGLQPETSYSLTVAITTKGDGARSKPKLVSTTGAVPKPLV 830
 Db 561 ADAQWEFDDTTEHDMIIISGLQPETSYSLTVAITTKGDGARSKPKLVSTTGAVPKPLV 611
 Qy 831 INHTQNTALIQHPPVDTFGPLOGVRLKPKGRKDMPLTTLFSEKEDHTATDIHKAS 890
 Db 612 INHTQNTALIQHPPVDTFGPLOGVRLKPKGRKDMPLTTLFSEKEDHTATDIHKAS 670
 Qy 891 YVFLSARNKVGFEEMVKEISIPPEVPTGFPONLHSEGTSTSVLSWQPPVLAERNGI 950
 Db 671 YVFLSARNKVGFEEMVKEISIPPEVPTGFPONLHSEGTSTSVLSWQPPVLAERNGI 730
 Qy 951 ITKYTLTYLRDINIPPLPMEQLIVPADTMTLGLKPDITYDVVKVRAHSTKGPSPSVQ 1010
 Db 731 ITKYTLTYLRDINIPPLPMEQLIVPADTMTLGLKPDITYDVVKVRAHSTKGPSPSVQ 790
 Qy 1011 FRTLPLVDQ-VFARNFHVKAVMKTSVLLSWEIPENYNSAMPFK-ILYDDGKQVEVDGRAT 1068
 Db 791 FRTLPLVDQSMFAKNFHVKAVMKTSVLLSWEIPENYNSAMPFK-ILYDDGKQVEVDGRAT 850
 Qy 1069 QKLVNMLKPKESYFVLTVNKGNSAGGLQHEVTAKTADVLRTKPAFTGKTNLDGMITVOL 1128
 Db 851 QKLVNMLKPKESYFVLTVNKGNSAGGLQHEVTAKTADVLRTKPAFTGKTNLDGMITVOL 910
 Qy 1129 PEVPANENIKGYIIIVPLKKSRSKFKIKPWESPEDELDLLEKESIRKRSIRYGREVEL 1188
 Db 911 PEVPANENIKGYIIIVPLKKSRSKFKIKPWESPEDELDLLEKESIRKRSIRYGREVEL 970
 Qy 1189 KPYIAAHFDVLPTFTLGDGDKHGGFTNKLOQSQEQYVFFVLAWEHAESKMYATSPYS 1248
 Db 971 KPYIAAHFDVLPTFTLGDGDKHGGFTNKLOQSQEQYVFFVLAWEHAESKMYATSPYS 1030
 Qy 1249 PVVSMDLDPQITDEEBGLIWWVGPVLAUVFIICIVITAILLYKRAESDSRKSIINNK 1308
 Db 1031 PVVSMDLDPQITDEEBGLIWWVGPVLAUVFIICIVITAILLYKRAESDSRKSIINNK 1090
 Qy 1309 EIPSHHPTDPEVLRRLNFQTPGMASHPPPIPILEADHIERLKANDNLKFSQYESIDPGQ 1368
 Db 1091 EIPSHHPTDPEVLRRLNFQTPGMASHPPPIPILEADHIERLKANDNLKFSQYESIDPGQ 1150
 Qy 1369 QFTWEHNSLVNPKPNRYANVIAVDHRSVRLSAIEGIPGSDYNNANVIGYRKQNAVIA 1428
 Db 1151 QFTWEHNSLVNPKPNRYANVIAVDHRSVRLSAIEGIPGSDYNNANVIGYRKQNAVIA 1210
 Qy 1429 QGSLPETFGDFWMIWEQSRATVMMTKLEERSVKCDQYWPGRGTGTHGLVQVTLDDTV 1488
 Db 1211 QGSLPETFGDFWMIWEQSRATVMMTKLEERSVKCDQYWPGRGTGTHGLVQVTLDDTV 1269
 Qy 1489 ELATYCVRTFALYKNGSSEKREVRQFQFTAWPDHGVPEHPTPPFLAFLRRVKTNPPDAGP 1548

Db 1066 PHTYFNVLNRRGSLGLQQTARTAFNMLSKPSPVAPKPDNDGFIIVVLPDQSQPVT 1125
Qy 1137 IKGYIIIVPLUKSR-GKFIKPWSPDEMDLLELLKISR-KRRSIRYGREVEL-KPYIA 1193
Db 1126 VONYFIWVPLKSRGQFPVLLGSPEDMDLEELIQDLSRLQQRVHRHSQLEVPFYIA 1185
Qy 1194 AHFDVLPTFTLGDGKHGGTNTKQSGQYVFPVLVAWEHAKMYATSPSDPVVSM 1253
Db 1186 ARFSILFAVHPGNKQKQGGSDNRGLBPHGRYVLPVLAVLQKNB-PTFAASPFSDPFDLD 1244
Qy 1254 DLDQPTIDDEGLIWWVGPVLAVFIIICIVIAILLYKR-----KRAESDSKSSIPNNKE 1309
Db 1245 NPDPQPIVDGEGLIWIGPVLAVFIIICIVIAILLYKPKDSRKSEPTKCLLNAD 1304
Qy 1310 IPSHPTDPVBLRLNFQTPGMASHPTPIPIELADHIERKANLNLPKQSEYISIDPQQ 1369
Db 1305 LAPHHPKDPVEMRRINFQTPGMLSHPTPIPTDMAEHMERLKANDSLKLSQSEYISIDPQQ 1364
Qy 1370 FTWEHSLNLEVNPKPNRYANVAYDHSRVLSAIEGIPGSDVYVNAVYIDYRKONAYATQ 1429
Db 1365 FTWEHSLNLEANKPNRYANVAYDHSRVILQPLEGIMGSDVYVNAVYIDYRKONAYATQ 1424
Qy 1430 GSLPETGDFWRMIWEQSRATVWMTKLEERSRVKCDQWPSRGTTGHLVQVTLDTVE 1489
Db 1425 GPLPETGDFWRVWVWEQSRATVWMTKLEERSKICDQWPNRGTYTYGFIQVTLDTME 1484
Qy 1490 LATVCVTRFALYKNGSSKREVRQFTAPDPHGVPEHTPFLAPLRVKTCNPPDAGPM 1549
Db 1485 LATFCVTRFSLKNGSSKREVRHFOFTAPDPHGVPEHTPFLAPLRVKTCNPPDAGPI 1544
Qy 1550 VVHCSAGVTRGCTIVIDAMLERIKHEKTVDIYGHVTLMAORNYMOTEDQYFIHDAL 1609
Db 1545 VVHCSAGVTRGCTIVIDAMLERIKHEKTVDIYGHVTLMAORNYMOTEDQYFIHEAL 1604
Qy 1610 LEAVTCGNTVEPARNLYAYIOKLTQIETGENVNTGMELEFKRLASSKAHTSRFISANLPCN 1669
Db 1605 LEAVSCGNTVEPARSLTYIYIQLAQVEGHEVNTGMELEFKRLASSKAHTSRFITASLPCN 1664
Qy 1670 KFKNRLNIMPYESTRVLQPIRGVSGDYINASFDIGYRQOKAYIATQGPLAETBDFW 1729
Db 1665 KFKNRLNILPYESSRVLQPIRGVSGDYINASFDIGYRQOKAYIATQGPLAETBDFW 1724
Qy 1730 RMLHEHNSIIVMLTKLEMREKCHOYWPERSARYQYFVVDPAEYNNMPOYILREPKV 1789
Db 1725 RALMENNSTIIVMLTKLEMREKCHOYWPERSARYQYFVVDPAEYNNMPOYILREPKV 1784
Qy 1790 TDARDGQSRTRVQFOFTDWPQGVPKSGEGFDIFGQVHKTKQFGDGPISVHCSAGVG 1849
Db 1785 TDARDGQSRTRVQFOFTDWPQGVPKSGEGFDIFGQVHKTKQFGDGPISVHCSAGVG 1844
Qy 1850 RTGVFTILSVILERNRYEGVDIFQTVKMLRTQRPAMVQTEDEQYFVSRAALEYLGSPDH 1909
Db 1845 RTGVFTILSVILERNRYEGVDIFQTVKMLRTQRPAMVQTEDEQYFVSRAALEYLGSPDH 1904
Qy 1910 YAT 1912
Db 1905 YAT 1907

RESULT 5
TDHULK
leukocyte antigen-related protein precursor - human
N:Alternate names: leukocyte common antigen homolog
N:Contains: protein-tyrosine-phosphatase (EC 3.1.3.48)
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004
C:Accession: S03841; J00051
R:Streuli, M.; Krueger, N.X.; Hall, L.R.; Schlossman, S.F.; Saito, H.
J. Exp. Med. 168, 1523-1530, 1988
A:Title: A new member of the immunoglobulin superfamily that has a cytoplasmic region ho
A:Reference number: J00051; MUID:89035978; PMID:2972792
A:Accession: S03841

A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1897 <STR>
A/Cross-references: UNIPROT:P10586; EMBL:Y00815; NID:G34266; PIDN:CAA68754.1; PID:G34267
C/Genetics:
A:Gene: GDB:PTPRF; LAR
A/Cross-references: GDB:120138; OMIM:179590
A:Map position: lp34-lp34
C/Superfamily: leukocyte antigen-related protein; fibronectin type III repeat homology;
ogy
C/Keywords: glycoprotein; phosphoprotein; phosphoric monoester hydrolase; transmembrane
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-1897/Product: leukocyte antigen-related protein #status predicted <MAT>
F;17-1250/Domain: extracellular #status predicted <EXT>
F;37-99/Domain: immunoglobulin homology <IMM1>
F;139-199/Domain: immunoglobulin homology <IMM2>
F;236-290/Domain: immunoglobulin homology <IMM3>
F;308-390/Domain: fibronectin type III repeat homology <FN3A>
F;403-489/Domain: fibronectin type III repeat homology <FN3B>
F;501-583/Domain: fibronectin type III repeat homology <FN3C>
F;586-585/Domain: fibronectin type III repeat homology <FN3D>
F;698-798/Domain: fibronectin type III repeat homology #status atypical <FN3E>
F;810-893/Domain: fibronectin type III repeat homology <FN3F>
F;905-989/Domain: fibronectin type III repeat homology <FN3G>
F;1001-1078/Domain: fibronectin type III repeat homology <FN3H>
F;1251-1274/Domain: transmembrane #status predicted <TM>
F;1275-1897/Domain: intracellular #status predicted <INT>
F;1285-1897/Domain: leukocyte common antigen cytosolic domain homology <LAC>
F;1365-1586/Domain: protein-tyrosine-phosphatase homology <PTP1>
F;1654-1877/Domain: protein-tyrosine-phosphatase homology <PTP2>
F;44-97,146-197,243-288/Diulfide bonds: #status predicted
F;107,240,285,711,956/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;1338/Active site: Cys (phosphocysteine intermediate) #status predicted
F;1544/Binding site: substrate phosphate (Arg) #status predicted
F;1829/Active site: Cys (phosphocysteine intermediate) #status predicted
F;1835/Binding site: substrate phosphate (Arg) #status predicted

Query Match 70.0% Score 7032; DB 1; Length 1897;
Best Local Similarity 69.8% Pred No. 0;
Matches 1338; Conservative 235; Mismatches 305; Indels 38; Gaps 14;
Qy 9 LLLLTFFLRDAETPRFTTRTPVDQTGVSQVAFICQATGDPKPIVMNKKGVSNQR 68
Db 8 LVMLGLVAGAGDSKPVFIKVPEDQTLGSGVAFVQATGEPKPRITWKKKKVSSOR 67
Qy 69 FEVIEFDGSGVLRIQPLTPRDEALIEYCVASNNVCEISVSTLTVLREDQIPRGPTI 128
Db 68 FEVIEFDGSGVLRIQPLTPRDEALIEYCVASNNVCEISVSTLTVLREDQIPRGPTI 127
Qy 129 DMGPQLKVVTRTRTATMLCAASGNPDPEITWFKDPLVDTSNNNGRIKQLRSESIGGTP 188
Db 128 DMGPQLKVVTRTRTATMLCAASGNPDPEITWFKDPLVDTSNNNGRIKQLRSESIGGTP 179
Qy 189 RGALQIEQSEESDQKYEYCVATNSAGTRYSAPANLYVRELREVRVPRPSIPTNHEIM 248
Db 180 -GALQIESEESDQKYEYCVATNSAGTRYSAPANLYVRELREVRVPRPSIPTNHEIM 234
Qy 249 PGGSVNTCTVAVGSPMPYVKKMLGAEDLTPEDDMPICGRNVLELNDVQCSANYTCVAMSTL 308
Db 235 PGGSVNTCTVAVGSPMPYVKKMLGAEDLTPEDDMPICGRNVLELNDVQCSANYTCVAMSTL 294
Qy 309 GVIEAIAQITVKALPKPPGTPVVTTESTATSTLTWDSGNPEPVSYYIIQHKPKNSEELYK 368
Db 295 GMIEATQVTVKALPKPPIDLVTTTATSTLTWDSGNSEPVYYIGIYRAAGTEGPFQ 354
Qy 369 EIDGVATTRYSVAGLSPSYDYEFVRVAVNNIGRPPSEPVLTQTSQAPSSAPSDVQARM 428
Db 355 EVDGVATTRYSIGLSLSPSEYAFRVLAVNSIGRPPSEAVRARTGEQAPSSPPRRVQARM 414
Qy 429 LSSSTIILVQWKEPEEPNGQIQGYRVYVYTMPTQVHNNMMKGNVADSDITITGNLVQKTY 488
Db 415 LLSASTMLVQWKEPEEPNGQIQGYRVYVYTPDSRRRPPNAWHKHNTDAGLLTIVGSLPGITY 474